a series of enzymes released from the tumor cells. At least one of these tumor-derived enzymes, urokinase plasminogen activator (uPA), converts plasminogen to plasmin, while a phosphoglycerate kinase from hypoxic tumor cells then reduces the plasmin so that it can be converted to angiostatin by one of several different metalloproteinases. Other types of tumors have since been reported to generate angiostatin, e.g., human prostate cancer. Prostate-specific antigen generates angiostatin-like fragments from plasminogen.

Furthermore, when murine fibrosarcoma cells were transfected with angiostatin, primary subcutaneous tumors formed. Their growth was slowed in proportion to increased levels of angiostatin production by the tumor cells. In these tumors the *total angiogenic output* of the primary tumor was decreased by transfected angiostatin, which opposed the activity of the tumor's secreted angiogenic promoter in a dose-dependent manner, but never completely counteracted it. The rate of tumor growth (expansion of tumor mass) was directly proportional to the total angiogenic output of the tumor, inversely proportional to angiostatin production and to tumor cell apoptosis, and virtually independent of tumor cell proliferation.

Murine hemangioendothelioma generated endostatin, a 20-kDa cleavage product of collagen XVIII. Human non-small cell lung carcinoma generated a 53-kDa cleavage product of antithrombin III (antiangiogenic ATIII). (This tumor does not metastasize, but the circulating angiogenesis inhibitor was detected because a subcutaneous tumor suppressed the growth of a second tumor at a remote site). All three of these proteins specifically inhibit endothelial cell proliferation and not other cell types. They have no effect on tumor cells per se. Endostatin inhibits tumor angiogenesis but not wound angiogenesis. It has no effect on pregnant mice nor on normal neonatal mice. Endostatin is present in *C. elegans* as a product of collagen XVIII and so may be at least 600 million years old on an evolutionary time scale. Other endogenous angiogenesis inhibitors are presented in Table 83-2.

Angiogenesis Inhibitors in Clinical Trial Angiogenesis inhibitors are in clinical trials in the United States for patients with cancer (Table 83-3). While a few endogenous antiangiogenic proteins have entered clinical trial [e.g., angiostatin, endostatin, and interleukin (IL) 12], these are more difficult to manufacture, and it will take some time before large quantities of these inhibitors become available for large numbers of patients. The majority of angiogenesis inhibitors currently in clinical trial for cancer are antibodies or small-molecular-weight synthetic molecules that inhibit specific targets along the angiogenic pathway.

BYSTANDER MOLECULES IN THE ANGIOGENIC PATHWAY

A variety of molecules in the angiogenic pathway are not strictly endothelial cell mitogens or suppressors but operate as modifiers, markers, or receptors of the angiogenic process. They include: (1) integrins $\alpha_v\beta_3$ and $\alpha_v\beta_5$, which are upregulated on proliferating endothelial cells and act as receptors for fragments of fibronectin and other matrix components; (2) ephrins, which specify arterial or venous development of capillary vessels; (3) cyclooxygenase 2 (COX 2), the production of which is stimulated by bFGF and which converts lipid precursors to prostaglandin E_2 (PGE_2), an angiogenic stimulator; (4) plasminogen activator inhibitor 1 (PAI-1), which counteracts the

upregulation ofuPA that is produced by growing capillaries, thus restricting proteolysis to a local event at the tip of angiogenic vessels; (5) AC133, a specific marker for circulating endothelial precursor cells, which arise from bone marrow; and (6) nitric oxide synthase, which generates nitric oxide, that induces vasodilation in the vascular bed, a possible prerequisite for sprout formation.

METASTASES ARE ANGIOGENESIS-DEPENDENT

Before tumors have become neovascularized in experimental animals, tumor cells rarely, if ever, shed into the circulation and metastases are essentially nonexistent. After the primary tumor becomes neovascularized, the number of tumor cells shed into the circulation increases in proportion to the increased neovascularization. Metastases that survive at a distant site must become angiogenic to be detected. Nonangiogenic metastases remain dormant at a microscopic size (<0.2 mm) indefinitely. Therefore, angiogenesis is required at both ends of the metastatic cascade.

Clinical patterns in which metastases first present may also be explained by angiogenic mechanisms. Cancer metastases are known to present in at least four different common clinical patterns and in one rare pattern. These clinical observations have previously been unrelated to each other but may be unified by angiogenic principles from tumor-bearing animals.

1. The patient whose metastases appear, sometimes explosively, a few months after surgical removal of a primary tumor (e.g., osteogenic sarcoma) may have lost a circulating angiogenesis inhibitor generated by the primary tumor. A model for this clinical presentation is the murine Lewis lung carcinoma that generates angiostatin.

2. Metastases that are not being suppressed by the primary tumor may already be present when the primary tumor is first diagnosed. The experimental model is a Lewis lung carcinoma subline that does not generate angiostatin.

3. The "unknown," or "occult," primary describes a pattern of metastases that present in the absence of a primary tumor or before it is located. In the relevant animal model lung metastases grow so rapidly that they suppress the primary tumor. However, it has not yet been determined whether a circulating angiogenic inhibitor is generated by the metastases.

4. If metastases do not appear until years after surgical removal of the primary tumor, the patient may harbor dormant microscopic metastases that are not angiogenic. Those that eventually switch to the angiogenic phenotype can grow to detectable metastases. An example is the node-negative breast cancer patient who develops lung metastases 10 to 15 years after resection of the primary tumor. An animal model that mimics this pattern has been developed. Surgical removal of a B-16 melanoma from a syngeneic mouse leaves numerous viable lung metastases that are not angiogenic and do not expand beyond 0.1 to 0.2 mm diameter; they remain dormant for the life of the animal. They also remain viable, as evidenced by the fact that trauma to the lung or transplantation of a small piece of lung to the subcutaneous tissue of another mouse quickly generates a lethal tumor in both cases.

5. After surgical removal of a renal cell carcinoma, metastases will sometimes regress completely or partly. While this is an uncommon clinical pattern, V2 carcinoma in the rabbit most closely resembles it. Removal of a primary tumor in the leg is followed by regression of metastases. This does not appear to be an immune reaction because fresh tumor grows successfully in the same rabbit. One explanation is that the metastases were dependent upon high production of a circulating angiogenic stimulator, such as bFGFfrom the primary tumor. High plasma levels of bFGF have been found to correlate with high mortality in human renal cancer.

These clinical patterns are summarized in Fig. 83-4 as a unifying guide for clinicians. The hypothesis that these clinical patterns are linked by angiogenic mechanisms requires additional confirmation in the laboratory, but it provides a direction for further research.

LEUKEMIA IS ANGIOGENESIS-DEPENDENT

Leukemia was assumed not to be angiogenic because it had been thought of as "liquid tumor." However, when bone marrow biopsies from children with newly diagnosed acute lymphoblastic leukemia were stained with an antibody to von Willebrand factor to highlight vascular endothelium, microvessel density was increased six- to sevenfold when compared to children with "control" bone marrow biopsies taken at the time of diagnosis of a solid tumor. Confocal microscopy further revealed that new microvessels in leukemic bone marrow were surrounded by a perivascular cuff of tumor cells like solid tumors.bFGF in the urine of the leukemic children was approximately sevenfold higher than in controls. Acute myeloid leukemia and chronic myeloid leukemia in adults are also associated with intense bone marrow neovascularization. Cellular levels of the angiogenic factorVEGF are significantly increased in acute myeloid leukemia and provide a prognostic indicator of outcome. The close physical configuration of microvessels and bone marrow cells may facilitate a two-way paracrine pathway between vascular endothelial cells that produce mitogens for bone marrow cells [such as granulocyte colony stimulating factor (G-CSF)] and bone marrow cells that produce bFGF, a mitogen for endothelial cells.

Further work suggests that leukemia growth is dependent on angiogenesis. Murine leukemias can be suppressed or eradicated and survival of leukemia-bearing mice prolonged when they are treated with antiangiogenic therapy. The mice were treated either by systemic administration of endostatin or by chemotherapy administered on an "antiangiogenic" schedule. A conventional schedule of chemotherapy at a maximum tolerated dose was ineffective (see below).

Certain patients with multiple myeloma refractory to all conventional therapy have undergone successful remission when treated with thalidomide. This has initiated a debate about whether the beneficial effect of thalidomide is due to its antiangiogenic activity or to some other property such as its weak capacity to inhibit tumor necrosis factor (TNF-a) activity. This question arose because microvessel density did not decrease significantly in parallel with improvement of the disease. However, this result is not unexpected. In animal tumors that undergo steady regression in tumor volume as a result of antiangiogenic therapy, microvessel density (microvessels per square millimeter) may in some cases remain constant even as tumor volume is reduced by

one-half, because capillary dropout and tumor cell dropout are going on at a near constant ratio. Furthermore, other effects of antiangiogenic therapy, such as the reduction of plasma leakage from tumor vessels, would not be revealed by microvessel density. Further, thalidomide is among the weakest inhibitors of TNF-a. Four other compounds that inhibit TNF-a more potently than thalidomide have no antiangiogenic effect. Pentoxifylline inhibits TNF-aat a similar potency as thalidomide, yet thalidomide inhibits cornea angiogenesis and pentoxifylline does not. Ibuprofen *increases* TNF-a in the serum of mice by twofold, yet it inhibits angiogenesis. Finally dexamethasone is a more potent inhibitor of TNF-a than thalidomide, but dexamethasone does not inhibit angiogenesis or does so only weakly.

ANTIANGIOGENIC THERAPY CIRCUMVENTS ACQUIRED DRUG RESISTANCE

The emergence of drug-resistant tumor cells is a major problem accompanying almost all chemotherapy. Conventional cytotoxic chemotherapy targets the cancer cell, and it is the genetic instability and high mutation rate of these cells that are responsible in part for acquired drug resistance. However, vascular endothelial cells are genetically stable and have a low mutation rate, like bone marrow cells. Bone marrow cells do not appear to develop drug resistance against conventional chemotherapy. Thus, it is possible that tumor vessels will also maintain sensitivity to anti-angiogenic therapy.

MICROVESSEL DENSITY IS A USEFUL PROGNOSTIC INDICATOR

Neovascularization in human brain tumors correlates directly with tumor grade. Tumor vascularity in cutaneous melanoma also influences prognosis. Microvessel density is an independent prognostic indicator for human breast cancer. In fact, the majority of reports (52 different studies) confirm that microvessel density is a powerful and often an independent prognostic indicator for a variety of different human cancers. However, at least seven other reports fail to show the prognostic value of microvessel density. Some of these negative reports may be methodologic problems. Others may represent the co-existence of angiogenesis inhibitors and stimulators that cannot easily be measured in a tumor. Summaries of all published reports are given in (Tables 83-4,83-5,83-6,83-7,83-8, and83-9). The best prognostic information from a histologic microsection of a tumor is obtained when the highest area of microvessel density ("hot spot") is quantified. These areas may contain the most angiogenic tumor cells, which have the highest chance of becoming an angiogenic metastasis.

Despite its usefulness as a*prognostic* marker, quantification of microvessel density is not necessarily a useful *surrogate* marker for efficacy of antiangiogenic therapy. Although it is currently employed in a few early clinical trials of antiangiogenic therapy, experimental studies suggest that it will be of little value to predict efficacy of an angiogenesis inhibitor. Microvessel density is determined mainly by intercapillary distance, itself governed by the cuff thickness of tumor cells surrounding a microvessel. In experimental animals, microvessel density may remain constant as a tumor is shrinking under antiangiogenic therapy. Residual tumor cells can form cuffs around remaining vessels, so that the vessel density may not change significantly. Microvessel density also may not distinguish between benign and malignant tumors. The microvessel density in normal pituitary tissue is higher than in a pituitary adenoma, which is higher than in a pituitary carcinoma. In the normal pituitary gland, the

perivascular cuff is one to two cell layers. However, in the adenoma the perivascular cuff is increased as tumor cells adapt to lower oxygen tensions. Cuff thickness is even greater in the carcinoma, thus giving the lowest microvessel density. However, for other tumors, such as breast carcinoma, microvessel density is significantly higher than in normal breast tissue. In certain animal tumors, angiogenic output exceeds the growth capacity of the tumor cells. Initial antiangiogenic therapy will lead to a reduction in microvessel density, which may then remain constant as vascular density comes into balance with tumor cell population, whether the tumor remains stable or regresses. Finally, if the first microvessel density is obtained from an open biopsy and a subsequent follow-up microvessel density is from a needle biopsy, the second density will be higher. This artifact is due to tissue compression by the needle biopsy.

CYTOTOXIC CHEMOTHERAPY MAY BE ANGIOGENESIS-DEPENDENT

Certain cytotoxic chemotherapeutic agents may depend in part for their anticancer activity on their ability to inhibit proliferating endothelial cells. Proliferating and migrating microvascular endothelial cells are exposed to chemotherapeutic drugs before tumor cells. However, the endothelial cells have a chance to recover during the traditional 2-3 week off-therapy period designed to allow recovery of bone marrow. Recovering endothelium can resupply residual tumor cells with new vessels and with paracrine factors necessary for tumor cell survival. Tumor recurrence requires resumption of chemotherapy, which itself can lead to emergence of drug-resistant clones of tumor cells (but not of endothelial cells). The convention of *maximum tolerated dose* (MTD) virtually forces the prolonged off-therapy schedule.

If the schedule of the chemotherapeutic agent is altered (increased number of lower drug doses) to apply maximum cytotoxic pressure to the endothelial cells in the tumor bed (i.e., an "antiangiogenic schedule" instead of a "conventional schedule"), even large tumors in mice may be permanently cured. In contrast, all mice on the conventional schedule ofMTDtherapy died with drug-resistant tumors. The antiangiogenic schedule of the chemotherapeutic drug was more frequent but was administered at a threefold lower total lower dose per day so that bone marrow was not depressed. Thus, efficacy of cytotoxic chemotherapy can be improved (in mice) by applying new logic to an old drug. The new logic is that antiangiogenic properties of certain conventional cytotoxic agents are not revealed unless the drugs are administered frequently. Frequent administration requires lower doses.

These results in mice may help to explain why some patients who are receiving long-term maintenance or even palliative chemotherapy continue to have stable disease beyond the time that tumor cells would have been expected to develop drug resistance. For example, long-term stable disease has been observed in a few patients with metastatic breast cancer who have been on weekly paclitaxel for several years. Paclitaxel has been reported to have antiangiogenic activity in addition to its anticancer activity. In the future, formal clinical trials with antiangiogenic schedules of chemotherapy should be tested.

Although "fractionated" radiotherapy (i.e., increased frequency of exposures at lower doses) was found empirically to be more effective and to cause fewer side effects than higher radiation doses more widely spaced, the biologic basis of this approach may be

due, in part, to the effect of ionizing radiation on endothelial cells. Conventional radiotherapy of tumor-bearing animals is greatly enhanced and with fewer side effects when the radiotherapy is administered in combination with subtherapeutic doses of angiostatin.

GUIDELINES FOR CLINICAL TRIALS OF ANGIOGENESIS INHIBITORS

New guidelines are required to examine the clinical effects of angiogenesis inhibitors because this class of drugs differs so markedly from cytotoxic chemotherapy. First, end-points for efficacy require different definitions. For cytotoxic therapy, lack of tumor regression is considered a failure, and an end-point of stable disease is little valued because it has never been shown to improve patient survival. In contrast, stable disease brought about by antiangiogenic therapy may be a favorable end-point of antiangiogenic therapy if it can be shown to improve the quality or duration of life. Experience with endostatin in early phase I clinical trials shows that patients with advanced metastatic disease refractory to all conventional therapy and who have had stable disease on endostatin for up to 6 months so far have pain relief, increased appetite, normal bone marrow function, and no side effects. A similar experience has been gained fromIFN-aadministered daily at low dose (3 million units) for malignancies dependent upon overexpression ofbFGF as their sole or main angiogenic protein (i.e., giant cell tumors of bone, angioblastoma). In fact, five of six of these patients who had failed all conventional therapy had complete regressions of their tumors by 1 to 3 years and are now off therapy and remain tumor free. This illustrates a second difference from cytotoxic therapy: antiangiogenic therapy takes longer to achieve stable disease and tumor regression is slower. It is analogous to the use of tamoxifen or to the treatment of other chronic diseases such as tuberculosis. Nevertheless, patients enjoy a high quality of life during antiangiogenic therapy.

Third, tumor progression during a clinical trial of cytotoxic chemotherapy is considered as a failure, and patients are often discontinued from the trial. With antiangiogenic therapy, some patients with rapidly advancing metastatic disease may show some tumor progression before stable disease is achieved. In the first clinical trials of tamoxifen, some patients were discontinued early in the trials because of tumor progression. After the term *tamoxifen flare* was invented, patients stayed on long-term tamoxifen therapy for several years. Of course, very rapid tumor progression during antiangiogenic therapy requires that the inhibitor be discontinued and the patient offered a different therapy.

Fourth, unlike cytotoxic chemotherapy, antiangiogenic therapy is more effective if it is administered frequently, without gaps, for a long period of time (e.g., like tamoxifen) and at a dose that has little or no toxicity. The term*MTD*is less useful for angiogenesis inhibitors.

Fifth, angiogenesis inhibitors can be used in combination with conventional chemotherapy, radiotherapy, immunotherapy, gene therapy, or other modalities, usually without increasing side effects. Clinical trials of angiogenesis inhibitors in combination with chemotherapy or radiotherapy are already under way.

CLINICAL SIGNS IN CANCER PATIENTS BASED ON ANGIOGENESIS

Certain clinical signs and symptoms from tumor neovascularization are associated with specific tumor types. For example, retinoblastomas in the posterior eye induce iris neovascularization in the anterior chamber. Some brain tumors induce angiogenesis in remote areas of the brain. Bone pain in metastatic prostate cancer may be related in part to neovascularization. A problem in the diagnosis of a primary bone tumor is that if the biopsy specimen contains only the neovascular response at the periphery of the tumor, it may be mistaken for granulation tissue or inflammation. Several cancer syndromes, such as inappropriate hormonal activity, hypercoagulation, and cachexia, are secondary to the presence of biologically active peptides released into the circulation from vascularized tumors. Therefore, an early therapeutic effect of antiangiogenic therapy could be increased appetite, weight gain, and disappearance of a cancer syndrome. The angiogenesis induced by cervical cancer may be observed by colposcopy; the appearance of telangiectasia, or "vascular spiders," in a mastectomy scar may herald local recurrence of tumor; color Doppler imaging can demonstrate neovascularization in breast cancer and other tumors; bladder carcinoma is detected by cystoscopy based, in part, on its neovascularization; and mammography may reveal the vascularized rim of a breast tumor. A wide range of radiologic signs of cancer are based on "enhancement" of lesions by radiopaque dyes sequestered transiently in the neovasculature of a tumor. Moreover, in some tumors large central areas cannot be penetrated by radiopaque dyes because of vascular compression, a situation that is unusual in prevascular tumors.

CLINICAL MISPERCEPTIONS ABOUT TUMOR ANGIOGENESIS

Because angiogenesis research is such a broad and rapidly moving field (at least 30 reports each week), certain misperceptions have emerged.

One misperception is that angiogenesis is synonymous with malignancy. The presence of angiogenesis does not distinguish between a benign and a malignant tumor. Benign adrenal adenomas are highly neovascularized but appear to lack the growth potential to take advantage of the new blood vessels they have induced. Angiogenesis may not be necessary for certain tumor cells that can grow as a flat sheet between membranes, e.g., gliomatosis in the meninges. Large tumors are thought to have "established" vessels that would be refractory to antiangiogenic therapy. A few feeder vessels, usually arteries, may be observed in the midst of a histologic cross-section of a tumor and could be considered as established. However, tumor cells depend on *thin-walled microvessels* for diffusion of nutrients, growth factors, and oxygen, and it is these vessels that continue to undergo high turnover rates even in a large, slowly growing or indolent tumors. These microvessels require the continuous presence of endothelial growth factors such as VEGF. Withdrawal or blockade of VEGF leads to endothelial cell apoptosis and regression of microvessels. In both animals and humans, very large tumors have regressed in response to antiangiogenic therapy, but a longer time of therapy is required. For example, a high-grade giant cell tumor (refractory to all conventional therapy) of >1 kg in the pelvis of a 17-year-old girl underwent 90% regression after 1 year of daily systemic therapy withIFN-a(3 million units). It is commonly stated that tumors "outgrow their blood supply." This is inaccurate; growing tumors can gradually *compress* their blood supply because of increasing interstitial pressure (discussed above). These areas of vascular compression become ischemic

but are not avascular. Necrosis may follow. Vessel compression also interferes with the optimal delivery of therapeutic agents. Paradoxically, antiangiogenic therapy can decrease ischemia, apparently because it decreases interstitial pressure.

Another misperception is that antiangiogenic therapy will be less effective against slowly growing tumors, because this is true for cytotoxic chemotherapeutic agents. In fact the opposite has been found in experimental animals. Slowly growing mouse tumors respond more effectively to angiogenesis inhibitors (TNP-470 or angiostatin) than do rapidly growing tumors. Rapidly growing tumors require higher doses of angiogenesis inhibitors to suppress their growth to the same extent as slowly growing tumors. While cytotoxic therapy is dependent on tumor cell cycle, antiangiogenic therapy is not. It is widely assumed that only "highly vascularized" tumors are susceptible to antiangiogenic therapy. This misperception comes from attempts to estimate the angiogenic output of a tumor from an angiogram or a gross tumor specimen. A large, dark, unstained area in an angiogram is usually due to nonfilling of compressed vessels. This is often misinterpreted as "avascular" tumor. However, at the microscopic level, histologic sections reveal high microvessel density. A large tumor observed at the operating table, such as a neurofibrosarcoma, may be a hard white mass and assumed to be "poorly vascularized," when in fact the histologic microsections show intense neovascularization.

SUMMARY: TWO CELLULAR TARGETS IN A TUMOR

An important lesson from angiogenesis research is to think about a tumor as containing two cell compartments that stimulate each other: the endothelial cell compartment and the tumor cell compartment. Anticancer therapy may be more efficacious if each compartment is treated by drugs that selectively target each cell type. The mutational rate is high in the tumor cell compartment and low in the endothelial cell compartment. This is the reason why it may be possible to employ antiangiogenic therapy for the long term, together with conventional chemotherapy or other therapies and subsequently in the postchemotherapy period.

DISEASES OF OCULAR NEOVASCULARIZATION

Pathologic angiogenesis is the most common cause of blindness worldwide. Pathologic neovascularization can occur in each compartment of the eye. For example, of>21 diseases that cause pathologic neovascularization in the cornea, contact lens wear, trauma, prior surgery, herpes simplex, and herpes zoster are the most frequently associated with pathologic neovascularization. Of ~37 diseases associated with iris neovascularization, central retinal vein occlusion, neovascular glaucoma, diabetes mellitus, and retinoblastoma are the most frequent. Of 14 diseases associated with retinal neovascularization, age-related macular degeneration, diabetes mellitus, retinopathy of prematurity, central retinal vein occlusion, branch retinal vein occlusion, and sickle cell disease are the most frequent. In western countries, age-related macular degeneration and diabetic retinopathy are the diseases of ocular neovascularization that affect the most patients. The large number of diseases listed above that are associated with ocular neovascularization and that directly or indirectly cause blindness illustrates how few effective therapies are currently available. This outline also reveals how few of these therapies can be administered systemically and how great is the opportunity to

employ antiangiogenic therapy in clinical trials to reduce the incidence of blindness from pathologic neovascularization.

AGE-RELATED MACULAR DEGENERATION

In age-related macular degeneration, angiogenesis occurs in the choroid. In the severe form of the disease, microhemorrhages from these new vessels lead to blindness. Approximately 1.7 million individuals in the United States suffer from the severe form, which is the leading cause of blindness in those³64 years. Laser therapy is less effective than in diabetic retinopathy. The angiogenic proteinVEGF is markedly elevated in macular degeneration and may be a major mediator of this disease. Of the five angiogenesis inhibitors currently in clinical trials for ocular neovascularization (Table 83-10), four are employed in the treatment of macular degeneration. One inhibitor is an antibody that neutralizes VEGF, and the other is a synthetic low-molecular-weight compound that targets a VEGF receptor.

DIABETIC RETINOPATHY

Diabetic retinopathy affects ~1.2 million of the estimated 14 million U.S. diabetic patients. It is the leading cause of blindness in persons between ages 25 and 64. Pathologic angiogenesis occurs in the retina, and new microvessels grow into the vitreous where they bleed and cause vitreous retraction. Laser therapy is more successful than in macular degeneration, but it is painful and causes gradual obliteration of the peripheral retina and loss of accompanying visual fields. Overexpression of VEGFmay also mediate diabetic retinopathy but appears to be induced by upregulation of HIF-1 secondary to hypoxia in the retina. An orally available protein kinase Cb inhibitor of VEGF is in phase III clinical trial for diabetic retinopathy. An early primary cause of the hypoxia may be adhesion of leukocytes to endothelium in retinal vessels (by upregulation of intercellular adhesion molecule 1 on retinal microvascular endothelium), leading to slow flow or periods of no flow.

RETINOPATHY OF PREMATURITY

At the time of birth, both the retina and its vascular supply are still growing. Blood vessels that supply the retina in the premature baby and in the newborn are exquisitely sensitive to changes in oxygen, a mechanism that guarantees an adequate blood supply to the growing retina. In newborn cats exposed to oxygen,VEGFlevels are downregulated and vascular growth in the retina is slowed or inhibited. However, the retina continues to grow. Subsequently, when the animal is returned to room air, the mismatch between the delayed vascularization and the steadily growing retina leads to relative hypoxia, which triggers a surge of VEGF and retinal neovascularization. This may cause retinal detachment and microhemorrhage. In the United States there are currently ~180,000 cases of retinopathy of prematurity, also called *retrolental fibroplasia*.

The increased understanding of the angiogenic mechanism of retinopathy of prematurity has led to clinical trials in which infants are weaned from oxygen to room air very slowly in order to prevent the rapid rise ofVEGF. A recent report showed that in newborn animals returned to room air after exposure to oxygen, pathologic neovascularization was completely prevented while normal vascular development continued if the animals

were treated with angiostatin for 5 days, beginning with the first day of exposure to room air. It is not yet clear whether systemic therapy of retinopathy of prematurity will be feasible in infants.

ENDOGENOUS ANGIOGENESIS INHIBITORS IN THE EYE

Normally the components of the eye that transmit light (cornea, aqueous, lens, and vitreous) are avascular. The maintenance of this avascular state is accomplished, in part, by the presence of potent inhibitors of angiogenesis. Pigment epithelium-derived factor (PEDF) is a 50-kDa serpin and is a potent angiogenesis inhibitor that is produced by retinal cells. The amount of inhibitory PEDF produced by retinal cells is directly correlated with oxygen concentrations, suggesting that its loss plays a permissive role in ischemia-driven retinal neovascularization. In other words, when oxygen is decreased PEDF is decreased and VEGF is increased. Both changes facilitate neovascularization. PEDF has the unique characteristic of inhibiting endothelial migration toward a wide variety of angiogenic inducers tested. It may be the predominant angiogenesis inhibitor in the eye. When neutralizing antibody to PEDF (but not preimmune sera) is injected into the cornea, it becomes neovascularized. PEDF has also been found in tumors.

Thrombospondin-1 has also been found in ocular tissues such as cornea and in the retina. During hypoxia-driven retinal angiogenesis in newborn mice (returned to room air after oxygen exposure), a threefold increase in expression of thrombospondin-1 was seen corresponding to peak neovascularization and peakVEGFexpression. The increased thrombospondin-1 expression during ischemia-induced angiogenesis appears to be mediated by VEGF. This suggests that thrombospondin functions in a negative-feedback system to protect the eye against surges of VEGF. It is interesting that the same balance of positive and negative regulators of angiogenesis originally found in tumors operates in normal tissues and that a shift in the net balance of these regulators mediates pathologic angiogenesis as well as its return to the normal nonangiogenic state.

ANGIOGENESIS IN SKIN DISEASE

Many dermatologic diseases are associated with angiogenesis. A caveat is that, unlike neoplastic diseases, which are virtually all angiogenesis-dependent, not all nonneoplastic diseases that are angiogenic are also angiogenesis-dependent.

ANGIOGENESIS-ASSOCIATED VS. ANGIOGENESIS-DEPENDENT SKIN DISEASE

In certain diseases of the skin, angiogenesis may be an important side effect that facilitates healing or otherwise protects the host. Examples include ulcerations, delayed healing of wounds, and chronic infections, in which antiangiogenic therapy could be contraindicated. A few of the dermatologic diseases known to be angiogenesis-dependent are described.

Infantile Hemangiomas Hemangiomas consist of tumor-like clusters of proliferating capillaries. They occur in 1 out of 100 newborns and in 1 out of 4 premature infants, and by age 1 year are present in up to 10% of infants. In the first, or proliferating, stage, the lesions grow rapidly, reaching peak growth by ~4 months. By about 1 year they may

enter the involuting stage, where they stop growing, following which they regress over the next 3 to 5 years (the involuted stage) and then usually disappear. During the proliferating stage the endothelial cells overexpressbFGF,VEGF, and metalloproteinases, all of which appear in the urine at abnormally high levels. In normal skin, keratinocytes expressIFN-b, an angiogenesis inhibitor of similar strength as IFN-a. IFN-a or -b inhibit overexpression ofaFGF and bFGF. Glucocorticoids (prednisone, 5 mg/kg) are used as first-line therapy for hemangiomas that are destroying tissue, interfering with sight, or threatening life. A dramatic slowing and subsequent regression occur in ~30% of patients, but glucocorticoids fail in the remaining 70% (i.e., either no regression, but some slowing of the disease, or continued rapid growth of the lesions). The mechanism by which glucocorticoids act as antiangiogenic agents is not clear, but they do inhibit synthesis of metalloproteinases. When glucorticorticoids fail and the hemangioma is life-threatening, IFN-a is used at low dose, 3 million units/m^2daily subcutaneously for 8 to 12 months. While ~95% of hemangiomas regress spontaneously, 5% are sight- or life-threatening. Hemangiomas in the liver, heart, airway, or brain may be associated with a 50% mortality if untreated. IFN-a is antiangiogenic on the basis of its ability to inhibit overproduction of bFGF. Urinary levels of bFGF fall toward normal as hemangiomas regress, and bFGF levels can be used as a guide to dosing of IFN-a. While IFN-aaccelerates regression of hemangiomas in 85% of patients, it fails in the other 15% for unknown reasons. Many of the failures are Kaposi hemangioendotheliomas (KHE), a very aggressive form of hemangioma, often accompanied by platelet trapping and thrombocytopenia. IFN-a works well in only 50% of KHE. Regressions of hemangioma are slower with IFN-a than with glucocorticoids. In infants<1 year old, a side effect of IFN-a can be delayed walking. This occurs in ~4% of infants and can be detected early by spasticity of the lower limbs (*spastic diplegia*). It is reversible if IFN-a is discontinued; for this reason, all children on IFN-aare followed carefully by a neurologist.

A "cavernous hemangioma" is not a hemangioma but a venous malformation in which there is a dearth of smooth muscle in the wall of a large thin venous structure lined by endothelium. These never regress spontaneously, and neither glucorticoids norIFN-aare effective. Thus an adult with a cavernous hemangioma should not be treated with IFN-a.

Verruca Vulgaris Warts are caused by infection of keratinocytes in the skin by one of many subtypes of human papillomavirus (HPV). HPV contains two genes (E6/E7) that may increase angiogenesis. The E6 gene destabilizes the p53 tumor-suppressor gene, which upregulatesVEGF and downregulates thrombospondin-1, an angiogenesis inhibitor. The HPV E7 gene inactivates the tumor suppressor gene Rb. Antiangiogenic therapy may be beneficial in these lesions, which are usually highly neovascularized.

Psoriasis Psoriasis is a proliferative disorder of epidermis accompanied by increased vascularity in the dermis in the form of elongated and widened dermal capillaries. The disease is T lymphocyte mediated. Psoriatic lesions are angiogenic. The major angiogenic mediator in psoriasis appears to beVEGF, which is upregulated, as are its receptors. In patients with psoriasis, the increased vascularity induced by VEGF may act as a conduit for delivery of T lymphocytes to the epidermal target. VEGF itself may facilitate T lymphocyte targeting. When VEGF is overexpressed in a tumor vascular bed, leukocyte rolling and adhesion are enhanced.

Up to 5 million Americans have psoriasis, but ~500,000 have a severe form that requires long-term therapy, such as with methotrexate for several years. Both glucocorticoids and retinoids are weak angiogenesis inhibitors, and they may be suppressing the angiogenic component as well as the infiltration of immune cells. More potent angiogenesis inhibitors may be useful.

Basal Cell and Squamous Cell Carcinomas Both of these skin malignancies are highly angiogenic and follow the rules for other angiogenic-dependent tumors. Inactivation of the p53 tumor-suppressor gene (in part by ultraviolet light-induced mutations) is thought to be an early event in tumorigenesis, and its inactivation downregulates the angiogenesis inhibitor thrombospondin-1 and upregulatesVEGFexpression. Furthermore, the normal expression ofIFN-bin keratinocytes is markedly decreased, permitting upregulation of the angiogenic stimulatorbFGF. These lesions comprise >90% of the ~700,000 skin cancers that are treated each year in the United States.

Cutaneous Melanoma Melanoma in the skin begins in a radial or horizontal growth phase, which usually does not exceed a thickness of 0.75 mm. This stage is not neovascularized or is poorly neovascularized. It is analogous to the avascular phase of early in situ carcinoma. In the vertical growth phase, there are increased neovascularization, increased proliferation, and increased thickness of tumor beyond 0.75 mm, and intensity of vascularization correlates directly with increased metastatic risk and mortality. Progression of melanoma often begins with inactivation of the tumor-suppressor gene p16 and is followed later by expression of$\alpha_v\beta_3$integrin and by expression ofVEGFreceptors on the melanoma cells. *Ras* mutations that upregulate VEGF expression emerge later and may be followed by expression of the angiogenic proteins,IL-8, andbFGF. This sequential onset of expression of angiogenic proteins by a tumor cell is similar to progression of breast cancer. It is not clear if or how expression of$\alpha_v\beta_3$integrin as well as VEGF receptors on the melanoma cells themselves facilitates tumor growth, unless the VEGF is acting as an autocrine growth factor for the tumor cells, while at the same time acting as a paracrine stimulator of endothelial cells. A precedent for this mechanism has been reported for human pancreatic cancer. The angiogenesis inhibitor 2-methoxyestradiol currently in a phase I trial for breast cancer also showed efficacy against melanoma in animals.

Kaposi's Sarcoma (KS) This lesion, which acts like a malignancy in patients with AIDS, may instead be a chronic inflammatory reactive process that is highly angiogenic. The angiogenesis is driven mainly byVEGF and hepatocyte growth factor. The tat protein of the HIV virus also plays a role in the angiogenic pathway for KS, but this is still being elucidated. The origin of KS cells is also not clear, but they appear to arise from the vascular system, possibly from smooth-muscle cells or pericytes. Two different angiogenesis inhibitors in phase II trials, thalidomide and TNP-470, a synthetic analogue of fumagillin, have shown efficacy against KS. However, there are currently too few cases to make any general conclusions.

Neurofibromatosis These slow-growing benign skin tumors of Schwann cell origin can grow in other organs and become very large. Neurofibromas are very neovascularized and expressVEGF. This may be driven by overexpression of the *ras* oncogene. The gene for neurofibromin (NF1) is a negative regulator of *ras*, and mutation of this gene

increases *ras* expression. Because of their very slow growth rate and high angiogenic activity, neurofibromas illustrate a type of tumor for which long-term antiangiogenic therapy may be more effective than cytotoxic chemotherapy. They are rich in mast cells, which may enhance tumor angiogenesis by mobilization ofbFGF. If a patient with a neurofibroma had abnormally high plasma or urine levels of bFGF, dailyIFN-acould be used at a low dose of 3 million units/$m^2$for a prolonged period of 2 to 3 years, with slow regression as a goal.

Recessive Dystrophic Epidermolysis Bullosa This autosomal recessive disorder is characterized by subepidermal blistering, scarring, fusion of digits, and severe pruritus. Epidermis separates from dermis, in part due to a loss of collagen VII, which participates in the anchoring of these two cellular layers. Aggressive cutaneous squamous cell carcinoma, which emerges from this lesion, is the most common cause of death. Very high levels ofbFGFhave been found in the urine of these patients but not in patients with other blistering disorders. The source of bFGF could be its mobilization from heparan sulfate proteoglycans in the defective epidermal-dermal junction. bFGF not only stimulates angiogenesis directly but also stimulates the production ofCOX 2, which converts lipid precursors to prostaglandin E_2, another angiogenic stimulator. Keratinocyte growth is also stimulated by bFGF. Continuous keratinocyte proliferation may be a precursor to the development of squamous cell carcinoma. COX 2 inhibitors have antiangiogenic and antitumor activity in mice and may be useful in angiogenic diseases in man. Another antiangiogenic approach could be low-doseIFN-a, based on the same rationale for reducing high bFGF expression in life-threatening hemangiomas.

ANGIOGENESIS IN ARTHRITIS

The role of angiogenesis in rheumatoid arthritis and in other forms of arthritis can be most simply conceptualized as two phases: prevascular and vascular. The prevascular phase is analogous to an acute inflammatory state in which the synovium is invaded by inflammatory and immune cells, with macrophages, mast cells, and T cells predominating, among others. These cells may be the source of the angiogenic stimulators found in synovial fluid, which includeVEGF,bFGF,IL-8, and hepatocyte growth factor. Activated endothelial cells can also release hepatocyte growth factor. The growth of a neovascular pannus from the synovium begins the vascular phase of arthritis. The vascular pannus can invade and destroy cartilage, a process that is enhanced by the generation of enzymatic activity, mainly metalloproteinases, at the advancing front of new proliferating endothelium. This neovascular pannus overcomes endogenous angiogenesis inhibitors in the cartilage that normally protect it from vascular invasion and maintain its avascularity. These inhibitors include, among others,TIMPs 1, 2, 3 and 4 (ranging from 21 to 29 kDa); thrombospondin-1; and troponin I. Experimental evidence that arthritis is angiogenesis-dependent is based on suppression of rat adjuvant arthritis by an angiogenesis inhibitor, TNP-470 (a synthetic analogue of fumagillin).

This somewhat simplistic model does not do justice to the complexity of the angiogenic response in arthritis, which is beyond the scope of this chapter. Nevertheless, it provides a platform to think about antiangiogenic therapy of arthritis. In principle, inhibition of neovascularization in the joint should interrupt a conduit for continuous traffic of inflammatory cells into the joint and prevent destruction of cartilage. TheCOX

2inhibitors currently in wide use for arthritis have been found to be potent angiogenesis inhibitors capable of inhibiting tumor growth in mice. Other angiogenesis inhibitors currently in clinical trial for cancer may eventually also find use in antiarthritis therapy. An interesting potential candidate would be 2-methoxyestradiol.

ANGIOGENESIS IN GYNECOLOGIC DISEASE

Angiogenesis in the female reproductive tract is being actively studied because it is the principal example of physiologic angiogenesis. Angiogenesis in the ovarian follicle is driven mainly bybFGF andVEGFalthough other angiogenic regulatory molecules are being studied. However, while VEGF is known to be upregulated by estrogen, it is not clear whether endogenous angiogenesis inhibitors operate in combination with declining estrogen to turn off angiogenesis in the ovarian follicle. When angiogenesis is increased in any one follicle, it is suppressed in all other follicles. This is analogous to the suppression of angiogenesis in distant metastases by a primary tumor, but it is not known if the ovarian system utilizes similar endogenous inhibitors as have been discovered in various tumor systems.

A variety of diseases of the female reproductive tract are based on angiogenic processes. A few of these are mentioned here briefly to illustrate that similar molecules mediate angiogenesis in tumors and in gynecologic disease, although they may be regulated differently.

Endometriosis In this disease, endometrial glands or stroma are present outside the uterine cavity, e.g., in the ovaries, uterine ligaments, rectovaginal septum, and pelvic peritoneum. The foci of endometrium are usually under the control of the ovarian hormones and undergo cyclic menstrual changes with periodic bleeding, which is painful and may lead to fibrosis. At least one angiogenic protein,VEGF, is known to mediate the neovascularization in these lesions. VEGF is upregulated by increased estrogen and downregulated by withdrawal of estrogen. It has been suggested that endometriotic tissue may produce its own estrogen because it contains aromatase cytochrome P450, not found in normal endometrium. Approximately 780,000 women in the United States suffer from endometriosis. No clinical trials of angiogenesis inhibitors for endometriosis are currently under way, but this class of drugs holds promise as an additional treatment of endometriosis, perhaps on a monthly basis. At least three angiogenesis inhibitors are produced in the female reproductive system: 2-methoxyestradiol, proliferin-related protein, and a 16-kDa fragment of prolactin. It would be of interest to know if any of these would be therapeutic for endometriosis.

Other Pathology Angiogenesis may be increased in dysfunctional uterine bleeding such as breakthrough bleeding from contraceptives. The edema and ascites of the ovarian hyperstimulation syndrome is thought to be mediated by the ability ofVEGF to increase vascular permeability. Preeclampsia during pregnancy may be related to abnormal vascular remodeling, although it is not clear how the hypertension and cerebral edema associated with this disease are mediated by an endothelial cell product.

Carcinoma of the Ovary, Endometrium, and Cervix These common gynecologic tumors are all angiogenesis-dependent. As a result, they share certain characteristics

discussed under "Neoplastic Disease," above. Microvessel density in histologic sections provides independent prognostic indicators of metastatic risk and/or mortality.VEGF is a major angiogenic mediator in these tumors. The ascites in ovarian carcinoma contains concentrations of VEGF of up to 100 times higher than those in serum in the same patient. Endometrial carcinoma, which can be induced by long-term tamoxifen therapy, may operate through upregulation ofIL-8 in the endometrium. The potential value of angiogenesis inhibitors in the treatment of gynecologic tumors refractory to conventional therapy is suggested by two reports: (1) experimental ovarian cancer was inhibited by administration of angiostatin and endostatin, which acted synergistically; and (2) malignant ascites and growth of human ovarian cancer were inhibited in experimental animals by inhibiting a receptor for VEGF.

ANGIOGENESIS IN CARDIOVASCULAR DISEASE

Angiogenesis in the cardiovascular system occurs under three different conditions: (1) neovascularization in atherosclerotic plaques, (2) formation of collateral vessels to an area of ischemic myocardium or ischemic muscle in a limb, and (3) neovascularization at the edges of a myocardial infarction during its repair.

Angiogenesis in Atherosclerotic Plaques Angiogenesis occurs in atherosclerotic plaques, and hypoxia is thought to be a major stimulus. The mediators of angiogenesis found in most plaques arebFGF,VEGF, transforming growth factor b(TGF-b), andPDGF-BB. Smooth-muscle cells in plaques are a source of VEGF and PDGF-BB. Macrophages, mast cells, and T cells, which are also found to infiltrate plaques, can produce bFGF and TGF-b. However, the angiogenic factor(s) directly responsible for plaque angiogenesis and the sequential order in which these factors act during the evolution of a plaque have not been worked out. The new microvessels in a plaque can be the source of intraplaque microhemorrhage. Furthermore, the production of metalloproteinases at the advancing tips of new microvessels may contribute to plaque rupture.

The evidence that atherosclerotic plaques are angiogenesis-dependent. However, some supporting experimental evidence has been obtained in transgenic mice deficient in the gene for apoliprotein E (ApoE-/-). When these mice are fed a western diet containing 0.15% cholesterol, they develop atherosclerotic plaques in the aorta over 6 months. Early plaques <250 um thick are not neovascularized. Cells in the center of such a plaque would lie within the oxygen diffusion limit of oxygen arriving from the normal vasa vasorum or from the arterial lumen. However, intense neovascularization occurred as plaques enlarged to >250 um. When mice were treated during the development of a plaque with either of the angiogenesis inhibitors TNP-470 (a synthetic fumagillin analogue in phase II clinical trial) or endostatin (in phase I clinical trial), total plaque area was reduced by 70% and 85%, respectively. This finding has important clinical implications.

If long-term antiangiogenic therapy inhibits plaque growth or reduces plaque microhemorrhage or rupture, then it will be important to document this in cancer patients who are receiving angiogenesis inhibitors over a period of [3]1 to 2 years. Furthermore, if antiangiogenic therapy blocks plaque angiogenesis, plaque growth, bleeding, or rupture, would this obviate the need for coronary collateral vessels? If so, then this would

remove the theoretical concern that long-term antiangiogenic therapy for cancer might decrease collateral development. Another extenuating circumstance is that collateral vessels in general are thick-walled and coated with smooth muscle. They are less likely to undergo regression during exposure to an angiogenesis inhibitor than the thin-walled endothelial tubes, which are not covered or stabilized by smooth muscle in a tumor bed.

Therapeutic Angiogenesis in Ischemic Vascular Disease Experimental and clinical attempts to increase angiogenesis in ischemic tissues are very recent and have generally followed two strategies: injection into the ischemic tissue of angiogenic proteins (eitherVEGF or FGFs) or injection of genetic material that codes for these angiogenic stimulators. The animal data show that it is possible to increase the density of new blood vessels and flow to an ischemic area beyond what can be accomplished by hypoxia defense mechanisms in the body. It is not yet clear how durable the new vessels will be once they are induced in an ischemic tissue. We understand much more about stopping angiogenesis than starting it. However, once the techniques for therapeutic angiogenesis are further developed, the clinical need could be enormous.

FUTURE DIRECTIONS

Many other diseases are dominated by the angiogenic process. These include Crohn's disease, thyroiditis, benign prostatic hypertrophy, glomerulonephritis, ectopic bone formation, keloids, and others. However, they were not included in this chapter because the evidence that they are angiogenesis-dependent is not yet sufficiently compelling.

The diseases that were included *are* more clearly angiogenesis-dependent and serve to illustrate an important direction for the future. Oncologists, dermatologists, ophthalmologists, rheumatologists, gynecologists, and cardiologists are dealing with diseases that appear on the surface to be completely different from each other. Advances in therapy of these diseases are reported at different meetings and in different journals, and the specialists who treat them rarely go to each other's meetings or talk to each other. Nevertheless, all of these diseases are dominated by pathologic angiogenesis. The angiogenesis is driven by a small but similar set of molecules, which are regulated differently in each disease. Furthermore, a new class of drugs, the angiogenesis inhibitors, is becoming available and may permit improvements in therapy for many of these diseases. Thus, angiogenesis is a unifying process that has heuristic value across many medical specialties. The "angiogenesis-dependency" of many diseases, neoplastic and nonneoplastic, is of course not sufficiently quantitative to be called a theory, but it is similar to Stephen Wolfram's definition of a theory as "a compressed package of information, applicable to many cases."

(Bibliography omitted in Palm version)

Back to Table of Contents

84. PRINCIPLES OF CANCER TREATMENT - *Edward A. Sausville*, *Dan L. Longo*

The goal of cancer treatment is first to eradicate the cancer. If this primary goal cannot be accomplished, the goal of cancer treatment shifts to palliation, the amelioration of symptoms, and preservation of quality of life while striving to extend life. The dictum *primum non nocere* is *not* the guiding principle of cancer therapy. Every cancer treatment has the potential to cause harm, and treatment may be given that produces toxicity with no benefit. The therapeutic index of many interventions is quite narrow, and most treatments are given to the point of toxicity. The guiding principle of cancer treatment is *primum succerrere*, first hasten to help. Radical surgical procedures, large-field hyperfractionated radiation therapy, high-dose chemotherapy, and maximum tolerable doses of cytokines such as interleukin (IL) 2 are all used in certain settings where 100% of the patients will experience toxicity and side effects from the intervention, and only a fraction of the patients will experience benefit. One of the challenges of cancer treatment is to use the various treatment modalities alone and together in a fashion that maximizes the chances for patient benefit.

Cancer treatments are divided into four main groups: surgery, radiation therapy (including photodynamic therapy), chemotherapy (including hormonal therapy), and biologic therapy (including immunotherapy, differentiating agents, and agents targeting cancer cell biology). The modalities are often used in combination, and agents in one category can act by several mechanisms. For example, cancer chemotherapy agents can induce differentiation, and antibodies (a form of immunotherapy) can be used to deliver radiation therapy. Surgery and radiation therapy are considered local treatments, though their effects can influence the behavior of tumor at remote sites. Chemotherapy and biologic therapy are usually systemic treatments.

Cancer behaves in many ways as an organ that regulates its own growth. However, cancers have not set an appropriate limit on how much growth should be permitted. Normal organs and cancers share the property of having a population of cells in cycle and actively renewing and a population of cells not in cycle. In cancers, cells that are not dividing are heterogeneous; some have sustained too much genetic damage to replicate but have defects in their death pathways that permit their survival; some are starving for nutrients and oxygen; and some are reversibly out of cycle poised to be recruited back into cycle and expand if needed. Severely damaged and starving cells are unlikely to kill the patient. The problem is that the cells that are reversibly not in cycle are capable of replenishing tumor cells physically removed or damaged by radiation and chemotherapy.

Tumors follow a Gompertzian growth curve (Fig. 84-1); the growth fraction of a neoplasm starts at 100% with the first transformed cell and declines exponentially over time until by the time of diagnosis at a tumor burden of 1 to 5´10^9tumor cells, the growth fraction is usually 1 to 4%. Cancers are actually trying to limit their own growth but are not completely successful at doing so. The peak growth rate occurs before the tumor is detectable. Folkman has suggested that tumors restrict their growth by elaborating angiogenesis inhibitors (Chap. 83). Other cellular mechanisms to withdraw cells from the cell cycle probably exist as well. Several observations support the idea of autoregulation of tumor growth. Metastases can be observed to grow more rapidly than the primary tumor, consistent with the idea that an inhibitory factor slows the growth of

larger tumor masses. When a tumor recurs after surgery or chemotherapy, frequently its growth is accelerated and the growth fraction of the tumor is increased. This pattern is similar to that seen in regenerating organs. Partial resection of the liver results in the recruitment of cells into the cell cycle, and the resected liver volume is replaced. Similarly, chemotherapy-damaged bone marrow increases its growth to replace cells killed by chemotherapy. However, cancers do not recognize a limit on their expansion. Monoclonal gammopathy of uncertain significance may be an example of a clonal neoplasm with intrinsic features that stop its growth before a lethal tumor burden is reached. A fraction of patients with this disorder go on to develop fatal multiple myeloma, but probably this occurs because of the accumulation of additional genetic lesions. Elucidation of the mechanisms that regulate this organ-like behavior may provide additional clues to cancer control and treatment.

PRINCIPLES OF CANCER SURGERY

Surgery is used in cancer prevention, diagnosis, staging, treatment (for both localized and metastatic disease), palliation, and rehabilitation.

PREVENTION

Cancer can be prevented by surgery in people who have premalignant lesions resected (e.g., premalignant lesions of skin, colon, cervix) and in those who are at higher-than-normal risk of cancer from either an underlying disease (colectomy in those with pancolonic involvement with ulcerative colitis), the presence of genetic lesions (familial polyposis -- colectomy; multiple endocrine neoplasia type II -- thyroidectomy; familial breast or ovarian cancer -- mastectomy, oophorectomy), or a developmental anomaly (orchiectomy in those with an undescended testis). In some cases, prophylactic surgery is more radical than the surgical procedures used to treat the cancer after it develops. The assessment of risk involves many factors and should be undertaken with care before advising a patient to undergo a major procedure. For breast cancer prevention, many experts use a 20% risk of developing breast cancer over the next 5 years as a threshold. However, patient fears play a major role in defining candidates for cancer prevention surgery. Counseling and education may not be enough to allay the fears of someone who has lost close family members to a malignancy.

DIAGNOSIS

The ideal diagnostic procedure varies with the type of cancer, its anatomic location, and the medical condition of the patient. However, the underlying principle is to obtain as much tissue as safely possible. Tumors may be heterogeneous in appearance. Pathologists are better able to make the diagnosis when they have more tissue to examine. In addition to light-microscopic inspection of a tumor for pattern of growth, degree of cellular atypia, invasiveness, and morphologic features that aid in the differential diagnosis, sufficient tissue is of value in searching for genetic abnormalities and protein expression patterns that may aid in differential diagnosis or provide information about prognosis or likely response to treatment. Such testing requires that the tissue be handled properly (e.g., immunologic detection of proteins is more effective in fresh-frozen tissue rather than in formalin-fixed tissue); thus, coordination among the

surgeon, pathologist, and primary care physician is essential to ensure that the amount of information learned from the biopsy material is maximized.

These goals are best met by an *excisional biopsy* in which the entire tumor mass is removed with a small margin of normal tissue surrounding it. If an excisional biopsy cannot be performed, *incisional biopsy* is the procedure of second choice. A wedge of tissue is removed, and an effort is made to include the majority of the cross-sectional diameter of the tumor in the biopsy to minimize sampling error. When the diagnosis is being made through an endoscope or via fluoroscopy, it may be necessary to obtain a *core-needle biopsy* of the mass; considerably less tissue is obtained and the diagnosis may be less certain. However, this procedure often provides enough information to plan a definitive surgical procedure. Least reliable in diagnosis of primary cancer is *fine-needle aspiration*. This technique generally obtains only a suspension of cells from within a mass. This approach (with stereotactic guidance of the needle) is the procedure of choice in the diagnosis of brain tumors and may be useful in diagnosing thyroid nodules and in confirming persistent or recurrent disease in a patient with known cancer, but the procedure is overutilized in primary diagnosis. It would be preferable to perform a larger open operation to obtain more tissue in most sites. The biopsy techniques that involve cutting into tumor carry with them a risk of facilitating the spread of the tumor.

STAGING

As noted in Chap. 79, an important component of patient management is defining the extent of disease. Radiographic and other imaging tests can be helpful in defining the clinical stage; however, pathologic staging requires defining the extent of involvement by documenting the histologic presence of tumor in tissue biopsies obtained through a surgical procedure. Axillary lymph node sampling in breast cancer and lymph node sampling at laparotomy for lymphomas and testicular, colon, and other intraabdominal cancers provide crucial information for treatment planning and may determine the extent and nature of primary cancer treatment.

TREATMENT

Surgery is perhaps the most effective means of treating cancer. About 40% of cancer patients are cured today by surgery. Unfortunately, a large fraction of patients with solid tumors (perhaps 60%) have metastatic disease that is not accessible for removal. However, even when the disease is not curable by surgery alone, the removal of tumor can obtain important benefits, including local control of tumor, preservation of organ function, debulking that permits subsequent therapy to work better, and staging information on extent of involvement. Cancer surgery aiming for cure is usually planned to excise the tumor completely with an adequate margin of normal tissue (the margin varies with the tumor and the anatomy), touching the tumor as little as possible to prevent vascular and lymphatic spread, and minimizing operative risk. Extending the procedure to resect draining lymph nodes obtains prognostic information, but such resections alone generally do not improve survival.

Increasingly, laparoscopic approaches are being taken to the primary tumor. Lymph node spread may be assessed using the *sentinel node approach*, in which the first

draining lymph node a tumor would encounter is defined by injecting a dye at operation and resecting the first node to turn blue. The sentinel node evaluation is continuing to undergo clinical testing but appears to provide reliable information without the risks (lymphedema, lymphangiosarcoma) associated with resection of all the regional nodes. Advances in adjuvant chemotherapy and radiation therapy following surgery have permitted a substantial decrease in the extent of primary surgery necessary to obtain the best outcomes. Thus, lumpectomy with radiation therapy is as effective as modified radical mastectomy for breast cancer, and limb-sparing surgery followed by adjuvant radiation therapy and chemotherapy has replaced radical primary surgical procedures involving amputation and disarticulation for childhood rhabdomyosarcomas. More limited surgery is also being employed to spare organ function, as in larynx and bladder cancer. The magnitude of operations necessary to optimally control and cure cancer has also been diminished by technical advances; for example, the circular anastomotic stapler has allowed narrower (<2 cm) margins in colon cancer without compromise of local control rates, and many patients who would have had colostomies are able to maintain normal anatomy.

In some settings, e.g., bulky testicular cancer or stage III breast cancer, surgery is not the first treatment modality employed. After an initial diagnostic biopsy, chemotherapy and/or radiation therapy are delivered to reduce the size of the tumor and control clinically undetected metastatic disease, and such therapy is followed by a surgical procedure to remove residual masses. This is called *neoadjuvant therapy*. Because the sequence of treatment is critical to success and is different from the standard surgery-first approach, coordination among the surgical oncologist, radiation oncologist, and medical oncologist is crucial.

Surgery may be curative in a subset of patients with metastatic disease. Patients with lung metastases from osteosarcoma may be cured by resection of the lung lesions. In patients with colon cancer who have fewer than five liver metastases restricted to one lobe and no extrahepatic metastases, hepatic lobectomy may produce long-term disease-free survival in 25% of selected patients. Surgery can also be associated with systemic antitumor effects. In the setting of hormonally responsive tumors, oophorectomy and/or adrenalectomy may control estrogen production and orchiectomy may reduce androgen production, both with effects on metastatic tumor growth. If resection of the primary lesion takes place in the presence of metastases, any noted change in tumor behavior is most often acceleration of growth, perhaps based on the removal of a source of angiogenesis inhibitors and mass-related growth regulators in the tumor. However, on rare occasions (certain renal cancers), primary tumor resection is accompanied by regression of metastatic lesions. Similarly, splenectomy in some cases of lymphoma may be associated with regression of disease at remote sites. This phenomenon is attributed to the removal of a source of growth or angiogenic factors upon which the remote sites depend for growth.

PALLIATION

Surgery is employed in a number of ways for supportive care: insertion of central venous catheters, diagnostic evaluation of pulmonary infiltrates, control of pleural and pericardial effusions and ascites, caval interruption for recurrent pulmonary emboli, stabilization of cancer-weakened weight-bearing bones, and control of hemorrhage,

among others. Surgical bypass of gastrointestinal, urinary tract, or biliary tree obstruction can alleviate symptoms and prolong survival. Surgical procedures may provide relief of otherwise intractable pain or reverse neurologic dysfunction (cord decompression). Splenectomy may relieve symptoms and reverse hypersplenism. Intrathecal or intrahepatic therapy relies on surgical placement of appropriate infusion portals. Surgery may correct other treatment-related toxicities such as adhesions or strictures.

REHABILITATION

Surgical procedures are also valuable in restoring a cancer patient to full health. Orthopedic procedures may be necessary to assure proper ambulation. Breast reconstruction can make an enormous impact on the patient's perception of successful therapy. Plastic and reconstructive surgery can correct the effects of disfiguring primary treatment.

PRINCIPLES OF RADIATION THERAPY

PHYSICAL PROPERTIES AND BIOLOGIC EFFECTS

Radiation therapy is a physical form of treatment that damages any tissue in its path. Tumor cells seem somewhat more sensitive to the lethal effects of radiation than normal tissues primarily because of differences in ability to repair sublethal DNA and other damage. In the target tissue, radiation damages DNA (usually single strand breaks) and generates free radicals from cell water that are capable of damaging cell membranes, proteins, and organelles. Radiation damage is dependent on oxygen; hypoxic cells are more resistant. Augmentation of oxygen is the basis for radiation sensitization. Sulfhydryl compounds interfere with free radical generation and may act as radiation protectors. The challenge for radiation treatment planning is to deliver the radiation to the tumor volume with as little normal tissue in the field as possible.*Principles of radiation injury are discussed in Chap. 394.*

Therapeutic radiation is delivered in three ways: teletherapy with beams of radiation generated at a distance and aimed at the tumor within the patient, brachytherapy with encapsulated sources of radiation implanted directly into or adjacent to tumor tissues, and systemic therapy with radionuclides targeted in some fashion to a site of tumor. Teletherapy is the most commonly used form of radiation therapy.

Radiation from any source decreases in intensity as a function of the square of the distance from the source (inverse square law). Thus, if the radiation source is 5 cm above the skin surface and the tumor is 5 cm below the skin surface, the intensity of radiation in the tumor will be $5^2/10^2$, or 25% of the intensity at the skin. By contrast, if the radiation source is moved to 100 cm from the patient, the intensity of radiation in the tumor will be $100^2/105^2$, or 91% of the intensity at the skin. Teletherapy maintains intensity over a larger volume of target tissue by increasing the source-to-surface distance. In brachytherapy, the source-to-surface distance is small; thus, the effective treatment volume is small.

X-rays and *gamma rays* are the forms of radiation most commonly used to treat cancer.

They are both electromagnetic, nonparticulate waves that cause the ejection of an orbital electron when absorbed. This orbital electron ejection is called *ionization*. X-rays are generated by linear accelerators; gamma rays are generated from decay of atomic nuclei in radioisotopes such as cobalt and radium. These waves behave biologically as packets of energy, called *photons*. Particulate forms of radiation are also used in certain circumstances. Electron beams have a very low tissue penetrance and are used to treat skin conditions such as mycosis fungoides. Neutron beams may be somewhat more effective than x-rays in treating salivary gland tumors. However, aside from these specialized uses, particulate forms of radiation such as neutrons, protons, and negative p mesons, which should do more tissue damage because of their higher linear energy transfer (LET) and be less dependent on oxygen, have not yet found wide applicability to cancer treatment.

A number of parameters influence the damage done to tissue by radiation. Hypoxic cells are relatively resistant. Nondividing cells are more resistant than dividing cells. In addition to these biologic parameters, physical parameters of the radiation are also crucial. The *energy* of the radiation determines its ability to penetrate tissue. Low-energy orthovoltage beams (150 to 400 kV) scatter when they strike the body, much like light diffuses when it strikes particles in the air. Such beams result in more damage to adjacent normal tissues and less radiation delivered to the tumor. Megavoltage radiation (³1 MeV) has very low lateral scatter; this produces a skin-sparing effect, more homogeneous distribution of the radiation energy, and greater deposit of the energy in the tumor, or *target volume*. The tissues that the beam passes through to get to the tumor is called the *transit volume*. The maximum dose in the target volume is often the cause of complications to tissues in the transit volume, and the minimum dose in the target volume influences the likelihood of tumor recurrence. Dose homogeneity in the target volume is the goal.

Radiation is quantitated on the basis of the amount of radiation absorbed in the patient, not based upon the amount of radiation generated by the machine. A rad (radiation absorbed dose) is 100 ergs of energy per gram of tissue; a gray (Gy) is 100 rad. Radiation dose is measured by placing detectors at the body surface or calculating the dose based on radiating phantoms that resemble human form and substance. Radiation dose has three determinants: total absorbed dose, number of fractions, and time. A frequent error is to omit the number of fractions and the duration of treatment. This is analogous to saying that a runner completed a race in 20 s; without knowing how far he or she ran, the result is difficult to interpret. The time could be very good for a 200-m race or very poor for a 100-m race. Thus, a typical course of radiation therapy should be described as 4500 cGy delivered to a particular target (e.g., mediastinum) over 5 weeks in 180-cGy fractions. Most radiation treatment programs are delivered once a day, 5 days a week in 150- to 200-cGy fractions.

The killing of tumor cells in vivo by radiation is described in detail inChap. 394. Although radiation can interfere with many cellular processes, many experts feel that a cell must undergo a double-stranded DNA break from radiation in order to be killed. The factors that influence tumor cell killing include the D_0 of the tumor (the dose required to deliver an average of one lethal hit to all the cells in a population), the D_q of the tumor (the threshold dose -- a measure of the cell's ability to repair sublethal damage), hypoxia, tumor mass, growth fraction, and cell cycle time and phase (cells in late G_1 and S are

more resistant). Rate of clinical response is not predictive; some cells do not die after radiation exposure until they attempt to replicate.

Compounds that incorporate into DNA and alter its stereochemistry (e.g., halogenated pyrimidines, cisplatin) augment radiation effects. Hydroxyurea, another DNA synthesis inhibitor, also potentiates radiation effects. Compounds that deplete thiols (e.g., buthionine sulfoximine) can also augment radiation effects. Hypoxia is the main factor that interferes with radiation effects.

APPLICATION TO PATIENTS

Radiation therapy can be used alone or together with chemotherapy to produce cure of localized tumors and control of the primary site of disease in tumors that have disseminated. Therapy is planned based on the use of a simulator with the treatment field or fields designed to accommodate an individual patient's anatomic features. Individualized treatment planning employs lead shielding tailored to shape the field and limit the radiation exposure of normal tissue. Often the radiation is delivered from two or three different positions. Conformal three-dimensional treatment planning is permitting the delivery of higher doses of radiation to the target volume without increasing complications in the transit volume.

Radiation therapy is a component of curative therapy for a number of diseases including breast cancer, Hodgkin's disease, head and neck cancer, prostate cancer, and gynecologic cancers. Radiation therapy can also palliate disease symptoms in a variety of settings: relief of bone pain from metastatic disease, control of brain metastases, reversal of cord compression and superior vena caval obstruction, shrinkage of painful masses, and opening threatened airways. In high-risk settings, radiation therapy can prevent the development of leptomeningeal disease and brain metastases in acute leukemia and lung cancer.

Brachytherapy involves placing a sealed source of radiation into or adjacent to the tumor and withdrawing the radiation source after a period of time precisely calculated to deliver a chosen dose of radiation to the tumor. This approach is often used to treat brain tumors and cervical cancer. The difficulty with brachytherapy is the short range of radiation effects (the inverse square law) and the inability to shape the radiation to fit the target volume. Normal tissue may receive substantial exposure to the radiation, with attendant radiation enteritis or cystitis in cervix cancer or brain injury in brain tumors.

TOXICITY

Though radiation therapy is most often administered to a local region, systemic effects, including fatigue, anorexia, nausea, and vomiting, may develop related in part to the volume of tissue irradiated, dose fractionation, radiation fields, and individual susceptibility. Bone is among the most radioresistant organs, radiation effects being manifested mainly in children through premature fusion of the epiphyseal growth plate. By contrast, the male testis, female ovary, and bone marrow are the most sensitive organs. Any bone marrow in a radiation field will be eradicated by therapeutic irradiation. Organs with less need for cell renewal, such as heart, skeletal muscle, and nerves, are more resistant to radiation effects. In radiation-resistant organs, the vascular

endothelium is the most sensitive component. Organs with more self-renewal as a part of normal homeostasis, such as the hematopoietic system and mucosal lining of the intestinal tract, are more sensitive. Acute toxicities include mucositis, skin erythema (ulceration in severe cases), and bone marrow toxicity. Often these can be alleviated by interruption of treatment.

Chronic toxicities are more serious. Radiation of the head and neck region usually produces thyroid failure. Cataracts and retinal damage can lead to blindness. Salivary glands stop making saliva, which leads to dental carries and poor dentition. Taste and smell can be affected. Mediastinal irradiation leads to a threefold increased risk of *fatal* myocardial infarction. Other late vascular effects include chronic constrictive pericarditis, lung fibrosis, viscus stricture, spinal cord transection, and radiation enteritis. The most serious late toxicity is the development of second solid tumors in or adjacent to the radiation fields. Such tumors can develop in any organ or tissue and occur at a rate of about 1% per year beginning in the second decade after treatment. Some organs vary in susceptibility to radiation carcinogenesis. Women under age 30 experience a 100-fold or greater increase in the incidence of breast cancer after chest or mantle field radiation; women treated after age 30 have little or no increased risk of breast cancer. No data suggest that a threshold dose of therapeutic radiation exists below which the incidence of second cancers is decreased. High rates of second tumors have been documented in people who received as little as 1000 cGy.

RADIONUCLIDES AND RADIOIMMUNOTHERAPY

Nuclear medicine physicians or radiation oncologists may administer radionuclides with therapeutic effects. Iodine-131 is used to treat thyroid cancer as iodine is naturally taken up preferentially by the thyroid. It emits gamma rays that destroy the normal thyroid as well as the tumor. Strontium-89 and samarium-153 are two radionuclides that are preferentially taken up in bone, particularly sites of new bone formation. Both are capable of controlling bone metastases and the pain associated with them, but the dose-limiting toxicity is myelosuppression.

Monoclonal antibodies and other ligands can be attached to radioisotopes by conjugation (for nonmetal isotopes) or by chelation (for metal isotopes), and the targeting moiety can result in the accumulation of the radionuclide preferentially in tumor. Iodine-131-labeled anti-CD20 and yttrium-90-labeled anti-CD20 are active in B cell lymphoma, and other labeled antibodies are being evaluated. Thyroid uptake of labeled iodine is blocked by cold iodine. Dose-limiting toxicity is myelosuppression.

PHOTODYNAMIC THERAPY

Some chemical structures (porphyrins, phthalocyanines) are selectively taken up by cancer cells by mechanisms not fully defined. When light, usually delivered by laser, is shone on cells containing these compounds, free radicals are generated and the cells die. Hematoporphyrins and light are being used with increasing frequency to treat skin cancer; ovarian cancer; and cancers of the lung, colon, rectum, and esophagus. Palliation of recurrent locally advanced disease can sometimes be dramatic and last many months.

PRINCIPLES OF CHEMOTHERAPY

HISTORIC BACKGROUND

The treatment of patients with cancer using chemicals in the hope of causing regressions of established tumors or to slow the rate of tumor growth arose by analogy to the proposition of Ehrlich that bacteria could be killed selectively by compounds acting as "magic bullets." Candidate compounds that might have selectivity for cancer cells were suggested by the marrow-toxic effects of sulfur and nitrogen mustards and led, in the 1940s, to the first notable regressions of hematopoietic tumors following use of these compounds by Gilman and colleagues. As these compounds caused covalent modification of DNA, the structure of DNA was thereby identified as a potential target for drug design efforts. Biochemical studies demonstrating the requirement of growing tumor cells for precursors of nucleic acids led to nearly contemporaneous studies by Farber of folate analogues. The cure of patients with advanced choriocarcinoma by methotrexate in the 1950s provided further impetus to define the value of chemotherapeutic agents in many different tumor types. This resulted in efforts to understand unique metabolic requirements for biosynthesis of nucelic acids and led to the design, rational for the time, of compounds that might selectively interdict DNA synthesis in proliferating cancer cells. The capacity of hormonal manipulations including oophorectomy and orchiectomy to cause regressions of breast and prostate cancers, respectively, provided a rationale for efforts to interdict various aspects of hormone function in hormone-dependent tumors. The serendipitous finding that certain poisons derived from bacteria or plants could affect normal DNA or mitotic spindle function allowed completion of the classic armamentatrium of "cancer chemotherapy agents" with proven safety and efficacy in the treatment of certain cancers.

END-POINTS OF DRUG ACTION

Chemotherapy agents may be used for the treatment of active, clinically apparent cancer.Table 84-1A lists those tumors considered curable by conventionally available chemotherapeutic agents. Most commonly, chemotherapeutic agents are used to address metastatic cancers. If a tumor is localized to a single site, serious consideration of surgery or primary radiation therapy should be given, as these treatment modalities may be curative as local treatments. Chemotherapy may be employed after the failure of these modalities to eradicate a local tumor, or as part of multimodality approaches to offer primary treatment to a clinically localized tumor. In this event, it can allow *organ preservation* when given with radiation, as in larynx or other upper airway sites; or sensitize tumors to radiation when given, for example, to patients concurrently receiving radiation for lung or cervix cancer (Table 84-1B). Chemotherapy can be administered as an *adjuvant* to surgery (Table 84-1C) or radiation, a use that may have curative potential in breast, colon, or anorectal neoplasms. In this use, chemotherapy attempts to eliminate clinically unapparent tumor that may have already disseminated. Chemotherapy can be used in *conventional dose* regimens. In general, these doses produce reversible acute side effects primarily consisting of transient myelosuppression with or without gastrointestinal toxicity (nausea), which are readily managed. *High-dose* chemotherapy regimens are predicated on the observation that the concentration-effect curve for many anticancer agents is rather steep, and increased dose can produce markedly increased therapeutic effect, although at the cost of potentially life-threatening

complications that require intensive support, usually in the form of bone marrow or stem cell support from the patient (*autologous*) or from donors matched for histocompatibility loci (*allogeneic*). High-dose regimens nonetheless have definite curative potential in defined clinical settings (Table 84-1D).

Karnofsky was among the first to champion the evaluation of a chemotherapeutic agent's benefit by carefully quantitating its effect on tumor size and using these measurements to decide objectively the basis for further treatment of a particular patient or further clinical evaluation of a drug's potential. A partial response (PR) is defined conventionally as a decrease by at least 50% in a tumor's bi-dimensional area; a complete response (CR) connotes disappearance of all tumor; progression of disease signifies increase by>25% from baseline or best response; and "stable" disease fits into none of the above categories.

If cure is not possible, chemotherapy may be undertaken with the goal of palliating some aspect of the tumor's effect on the host. Common tumors that may be meaningfully addressed with palliative intent are listed inTable 84-1E. Usually tumor-related symptoms may manifest as pain, weight loss, or some local symptom related to the tumor's effect on normal structures. Patients treated with palliative intent should be aware of their diagnosis and the limitations of the proposed treatment, have access to suitable palliative strategies in the event that no treatment is elected, and have a suitable "performance status" [according to assessment algorithms such as the one developed by Karnofsky or by the Eastern Cooperative Oncology Group (ECOG)]. ECOG performance status 0 (PS0) patients are without symptoms; PS1 patients have mild symptoms not requiring treatment; PS2, symptoms requiring some treatment; PS3, disabling symptoms, but allowing ambulation for>50% of the day; PS4, ambulation<50% of the day. Only PS0 to PS2 patients are generally considered suitable for palliative (noncurative) treatment. If there is curative potential, even poor performance status patients may be treated, but their prognosis is usually inferior to those of good performance patients treated with similar regimens.

PATH FOR NEW DRUG DISCOVERY AND DEVELOPMENT

The usefulness of any drug is governed by the extent to which a given dose causes a useful result (therapeutic effect; in the case of anticancer agents, toxicity to tumor cells) as opposed to a toxic effect. The therapeutic index is the degree of separation between toxic and therapeutic doses. Really useful drugs have large therapeutic indices, and this usually occurs when the drug target is expressed in the disease-causing compartment as opposed to the normal compartment. Classically, selective toxicity of an agent for an organ is governed by the expression of an agent's target; or differential accumulation into or elimination from compartments where toxicity is experienced or ameliorated, respectively. Current antineoplastic agents have the unfortunate property that their targets are present in both normal and tumor tissues. In the main they therefore have relatively narrow therapeutic indices.

Agents with promise for the treatment of cancer have in the past been detected empirically through screening for antiproliferative effects in animal or human tumors in rodent hosts or through inhibition of tumor cells growing in tissue culture. An optimal schedule for demonstrating antitumor activity in animals is defined in further preclinical

studies, as is the optimal drug formulation for a given route and schedule. Safety testing in two species on an analogous schedule of administration defines the starting dose for a phase I trial in humans, where escalating doses of the drug are given until reversible toxicity is observed. *Dose-limiting toxicity* (DLT) defines a dose that conveys greater toxicity than would be acceptable in routine practice, allowing definition of a *maximal tolerated dose* (MTD). The occurrence of toxicity is correlated if possible with plasma drug concentrations. The MTD or a dose just lower than the MTD is usually the dose suitable for phase II trials, where a fixed dose is administered to a relatively homogeneous set of patients in an effort to define whether the drug causes regression of tumors. An "active" agent conventionally has partial response rates of at least 20 to 25% with reversible non-life-threatening side effects, and it may then be suitable for study in phase III trials to assess efficacy in comparison to standard or no therapy. Response is but the most immediate indicator of drug effect. To be clinically valuable, responses must translate into effects on *overall survival* or at least *time to progression* as important indicators of an ultimately useful drug. More recently, active efforts to quantitate effects of anticancer agents on *quality of life* as an important outcome are being developed. Cancer drug clinical trials conventionally use a toxicity grading scale where grade I toxicities do not require treatment; grade II often require symptomatic treatment but are not life-threatening; grade III toxicities are potentially life-threatening if untreated; grade IV toxicities are actually life-threatening; and grade V toxicities ultimately lead to patient death.

The process of cancer drug development is likely to evolve in significant ways in the near future as (1) the molecular analysis of human tumors defines more precisely the molecular targets that can be the focus of drug discovery efforts, and (2) clinical trials are undertaken only after means of assessing the behavior of the drug in relation to its target have been developed. The basis for optimism and anticipated change in clinical trials methodology extends from emerging understanding of the basis for cancer incidence and progression. Cancer arises from genetic lesions that cause an excess of cell growth or division, with inadequate cell death (Chap. 82). In addition, failure of cellular differentiation results in altered cellular position and capacity to proliferate while cut off from normal cell regulatory signals. An overall schema for understanding cancer progression can be seen inFig. 84-2. Normally, cells in a differentiated state are stimulated to enter the cell cycle from a quiescent state, or "G0," or continue after completion of a prior cell division cycle in response to environmental cues including growth factor and hormonal signals. Cells progress through G1 and enter S phase after passing through "checkpoints," which are biochemically regulated transition points, to assure that the genome is ready for replication. One important checkpoint is mediated by the p53 tumor-suppressor gene product, acting through its upregulation of the $p21^{WAF1}$inhibitor of cyclin-dependent kinase (CDK) function, acting on CDKs 4 or 6. These molecules can also be inhibited by the $p16^{INK4A}$and $p27^{KIP1}$CDK inhibitors and, in turn, are activated by cyclins of the D family (which appear during G1) and the proper sequence of regulatory phosphorylations. Activated CDKs 4 or 6 phosphorylate and thus inactivate the product of the retinoblastoma susceptibility gene, pRb, which in its nonphosphorylated state complexes with transcription factors of the E2F family. Phosphorylated pRb releases E2Fs, which activate genes important in completing DNA replication during S phase, progression through which is promoted by CDK2 acting in concert with cylins A and E. During G2, another checkpoint occurs, in which the cell assures the completion of correct DNA synthesis. Cells then progress into M phase

under the influence of CDK1 and cyclin B. Cells may then go on to a subsequent division cycle or enter into a quiescent, differentiated state.

Also shown inFig. 84-2 are the sites of action of protooncogenes, regulators of cellular proliferation that, in an active state, promote cell growth, and whose deregulation produces oncogenes, originally discovered as the genes encoded by tumor-forming viruses in animals. Oncogenes can be divided into two families: (1) those that act in the cytoplasm to disrupt normal growth factor-related signaling, including *ras*, *raf*, and the tyrosine kinases of the *src* and *erb*B or *sis* families; and (2) nuclear oncogenes, including *jun*, *fos*, *myc*, and *myb*, that act to alter transcriptional control of cassettes of genes. In contrast, tumor-suppressor genes, including p53 and pRb, act as cellular "brakes" whose normal function is to inhibit or prevent unregulated cellular growth. The capacity to divide indefinitely is provided by activation of *telomerase*, which allows continued replication of chromosomes by addressing the unique need of chromosome ends to be continually renewed to a proper length to allow normal mitosis. The capacity to invade and metastasize is conveyed by elaboration of *matrix metalloproteases* and *plasminogen activators* and the capacity to recruit host stromal cells at the site of invasion through tumor-induced *angiogenesis*.

As will become apparent below, currently used drugs for the treatment of cancer focus principally on the proximate biochemistry of nucleic acid and mitotic spindle structure or function. Drugs of the future may seek to replace lost function of tumor-suppressor genes; counter the action of activated oncogenes; influence the capacity of cells to die; prevent normal chromosomal end replication; actually infect cells with viruses designed to replicate in the milieu of the cancer but not the normal cell; cause differentiation of cells with exit from the cell cycle by activating the appropriate genes; and utilize immunologic strategies, including antibodies and engineered cells to be directed at novel proteins expressed on the surface of cancer cells.

BIOLOGIC BASIS FOR CANCER CHEMOTHERAPY

The classic view of how cancer chemotherapeutic agents cause regressions of tumors focused on models such as the L1210 murine leukemia system, where cancer cells grow exponentially after inoculation into the peritoneal cavity of an isogenic mouse. The interaction of drug with its biochemical target in the cancer cell was proposed to result in "unbalanced growth" that was not sustainable and therefore resulted in cell death, directly as a result of interacting with the drug's proximal target. Agents could be categorized (Fig. 84-3) as cell cycle-active, phase-specific (e.g., antimetabolites, purines, and pyrimidines in S phase; vinca alkaloids in M), and phase-nonspecific agents (e.g., alkylators, and antitumor antibiotics including the anthracyclines, actinomycin, and mitomycin), which can injure DNA at any phase of the cell cycle but appear to then block in G2 before cell division at a checkpoint in the cell cycle. Cells arrested at a checkpoint may repair DNA lesions. Checkpoints have been defined at the G1 to S transition, mediated by the tumor-suppressor gene p53 (giving rise to the characterization of p53 as a "guardian of the genome"); at the G2 to M transition, mediated by the *chk1* kinase influencing the function ofCDK1; and during M phase, to ensure the integrity of the mitotic spindle. The importance of the concept of checkpoints extends from the hypothesis that repair of chemotherapy-mediated damage can occur while cells are stopped at a checkpoint; therefore, manipulation of checkpoint function

emerges as an important basis of affecting resistance to chemotherapeutic agents.

Resistance to drugs was postulated to arise either from cells not being in the appropriate phase of the cell cycle or from decreased uptake, increased efflux, metabolism of the drug, or alteration of the target, e.g., by mutation or overexpression. Indeed, the *p170PGP* (p170 P-glycoprotein; *mdr* gene product) was recognized from experiments with cells growing in tissue culture as mediating the efflux of chemotherapeutic agents in resistant cells. Certain neoplasms, particularly hematopoietic tumors, have an adverse prognosis if they express high levels of p170PGP, and modulation of this protein's function has been attempted by a variety of strategies.

Combinations of agents were proposed to afford the opportunity to affect many different targets or portions of the cell cycle at once, particularly if the toxic effects for the host of the different components of the combination were distinct. Combinations of agents were actually more effective in animal model systems than single agents, particularly if the tumor cell inoculum was high. This thinking led to the design of "combination chemotherapy" regimens, where drugs acting by different mechanisms (e.g., an alkylating agent plus an antimetabolite plus a mitotic spindle blocker) were combined. Particular combinations were chosen to emphasize drugs whose individual toxicities to the host were, if possible, distinct.

This view of cancer drug action is grossly oversimplified. Most tumors do not grow in an exponential pattern but rather follow Gompertzian kinetics, where the rate of tumor growth decreases as tumor mass increases (Fig. 84-1). Thus, a tumor has quiescent, differentiated compartments; proliferating compartments; and both well-vascularized and necrotic regions. Also, cell death is itself now understood to be a closely regulated process. *Necrosis* refers to cell death induced, for example, by physical damage with the hallmarks of cell swelling and membrane disruption. *Apoptosis*, or programmed cell death, refers to a highly ordered process whereby cells respond to defined stimuli by dying, and it recapitulates the necessary cell death observed during the ontogeny of the organism. *Anoikis* refers to death of epithelial cells after removal from the normal milieu of substrate, particularly from cell-to-cell contact. Cancer chemotherapeutic agents can cause both necrosis and apoptosis. Apoptosis is characterized by chromatin condensation (giving rise to "apoptotic bodies"); cell shrinkage; and, in living animals, phagocytosis by surrounding stromal cells without evidence of inflammation. This process is regulated either by signal transduction systems that promote a cell's demise after a certain level of insult is achieved or in response to specific cell-surface receptors that mediate cell death signals. Modulation of apoptosis by manipulation of signal transduction pathways has emerged as a basis for understanding the actions of currently used drugs and designing new strategies to improve their use.

The current view envisions that the interaction of a chemotherapeutic drug with its target causes or is itself a signal that initiates a "cascade" of signaling steps to trigger an "execution phase" where proteases, nucleases, and endogenous regulators of the cell death pathway are activated. Effective cancer chemotherapeutic agents are efficient activators of apoptosis through signal transduction pathways (Fig. 84-4). For example, in the cytokine-mediated pathway, exogenous ligands such as the Fas ligand (FasL) bind to cell-surface receptors (CD95; Fas), or tumor necrosis factor (TNF) or its homologue

Apo2L binds to its cognate receptors and directly recruits accessory molecules to activate a protease cascade (utilizing members of the caspase family of *c*ysteine *asp*artyl prote*ases*), resulting in apoptosis. In a second pathway, growth factor deprivation elicits poorly defined signals that result in protease activation. Chemotherapeutic agents create molecular lesions (in DNA or cellular membranes) as a consequence of combining with their respective molecular targets. These lesions are sensed by a cellular "damage sensor," whose molecular nature is unclear, which leads to mitochondrial damage. Release of mitochondrial factors (e.g., APAF1, cytochrome c) promotes the activation of another set of caspases. Damage to the plasma membrane, e.g., from free radicals generated by certain chemotherapeutic agents, leads to activation of acid sphingomyelinase to release lipid components including ceramides, which then promote apoptosis through a variety of pathways including direct mitochondrial damage.

While apoptotic mechanisms are important in regulating cellular proliferation and the behavior of tumor cells in vitro, in vivo it is unclear whether all of the actions of chemotherapeutic agents to cause cell death can be attributed to apoptotic mechanisms. Loss of clonogenic survival (conventionally detecting the capacity of a few cells to survive) may predict clinical value more reliably than detection of apoptotic changes in the majority of tumor cells. However, changes in molecules that regulate apoptosis are clearly correlated with clinical outcomes (e.g., *bcl2* overexpression in certain lymphomas conveys poor prognosis; proapototic *bax* expression is associated with a better outcome in ovarian carcinoma). Further efforts to understand the relationship of cell death and cell survival mechanisms will be necessary.

CHEMOTHERAPEUTIC AGENTS USED FOR CANCER TREATMENT

Table 84-2 lists commonly used cancer chemotherapy agents and pertinent clinical aspects of their use. The drugs may be usefully grouped into three general categories: those affecting DNA, those affecting microtubules, and those acting at hormone-like receptors.

Direct DNA-Interactive Agents

Formation of covalent DNA adducts Alkylating agents as a class break down, either spontaneously or after normal organ or tumor cell metabolism, to reactive intermediates that covalently modify bases in DNA. This leads to cross-linkage of DNA strands or the appearance of breaks in DNA as a result of repair efforts. "Broken" or cross-linked DNA is intrinsically unable to complete normal replication or cell division; in addition, it is a potent activator of cell cycle checkpoints and signaling pathways that can activate apoptosis. As a class, alkylating agents share similar toxicities, including myelosuppression, alopecia, gonadal dysfunction, mucositis, and pulmonary fibrosis. They differ greatly in a spectrum of normal organ toxicities.

Nitrogen mustard (mechlorethamine) is the prototypic agent of this class, decomposing rapidly in aqueous solution to yield potentially a bifunctional carbonium ion. It must be administered shortly after preparation into a rapidly flowing intravenous line. It is powerful vesicant, and infiltration may be symptomatically ameliorated by infiltration of the affected site with 1/6 *M* thiosulfate. Even without infiltration, aseptic thrombophlebitis

is frequent. It can be used topically as a dilute solution in cutaneous lymphomas, with a notable incidence of hypersensitivity reactions. It causes moderate nausea after intravenous administration.

Cyclophosphamide is inactive unless metabolized by the liver to 4-hydroxyl-cyclophosphamide, which decomposes into alkylating species, as well as to chloroacetaldehyde and acrolein. The latter causes chemical cystitis, and therefore excellent hydration must be maintained while using cyclophosphamide. If severe, the cystitis may be effectively treated by mercaptoethanesulfonate (MESNA). Liver disease impairs drug activation. Sporadic interstitial pneumonitis leading to pulmonary fibrosis can accompany the use of cyclophosphamide, and high doses used in conditioning regimens for bone marrow transplant can cause cardiac dysfunction. Ifosfamide is a cyclophosphamide analogue also activated in the liver, but more slowly, and it requires mandatory coadministration of MESNA to prevent bladder injury. Central nervous system (CNS) effects, including somnolence, confusion, and psychosis, can follow ifosfamide use, and the incidence appears related to low body surface area or the presence of nephrectomy.

There are several less commonly used alkylating agents. Chlorambucil causes predictable myelosuppression, azospermia, nausea, and pulmonary side effects. Busulfan can cause profound myelosuppression, alopecia, and pulmonary toxicity but is relatively "lymphocyte sparing." Its routine use in treatment of chronic myeloid leukemia has been curtailed in favor of hydroxyurea or interferon (IFN), but it still is employed in marrow transplant preparation regimens. Melphalan shows variable oral bioavailability and undergoes extensive binding to albumin anda_1-acidic glycoprotein. Mucositis appears more prominently.

Nitrosoureas break down to carbamoylating species that not only cause a distinct pattern of DNA base pair-directed toxicity but also can covalently modify proteins. They share the feature of causing relatively delayed bone marrow toxicity, which can be cumulative and long-lasting. Streptozotocin is unique in that its glucose-like structure conveys specific toxicity to the islet cells of the pancreas (for whose derivative tumor types it is prominently indicated) as well as causing renal toxicity in the form of Fanconi's syndrome, including amino aciduria, glycosuria, and renal tubular acidosis. Methyl CCNU (lomustine) causes direct glomerular as well as tubular damage, cumulatively related to dose and time of exposure.

Procarbazine is metabolized in the liver and possibly in tumor cells to yield a variety of free radical and alkylating species. In addition to myelosuppression, it causes hypnotic and other CNS effects, including vivid nightmares. It can cause a disulfiram-like syndrome on ingestion of ethanol. Hexamethylmelamine and thiotepa can chemically give rise to alkylating species, although the nature of the DNA damage has not been well characterized in either case. Thiotepa can be used for intrathecal treatment of neoplastic meningitis. Dacarbazine (DTIC) is activated in the liver to yield the highly reactive methyl diazonium cation. It causes only modest myelosuppression from 21 to 25 days after a dose but causes prominent nausea on day 1.

Cisplatin was discovered fortuitously by observing that bacteria present in electrolysis solutions could not divide. Only the cis diamine configuration is active as an antitumor

agent. It is hypothesized that in the intracellular environment, a chloride is lost from each position, being replaced by a water molecule. The resulting positively charged species is an efficient bifunctional interactor with DNA, forming Pt-base cross-links. Cisplatin requires administration with adequate hydration, including forced diuresis with mannitol to prevent kidney damage; even with the use of hydration, gradual decrease in kidney function is common. Hypomagnesemia frequently attends cisplatin use and can lead to hypocalcemia and symptomatic tetany. Other common toxicities include neurotoxocity with "stocking and glove" sensorimotor neuropathy. Hearing loss occurs in 50% of patients treated with conventional doses. Cisplatin is intensely emetogenic, requiring prophylactic antinausea agents. Myelosuppression is less evident than with other alkylating agents. Chronic vascular toxicity is a more unusual toxic phenomena, including Raynaud's syndrome and coronary artery disease. In an effort to obviate these toxicities, carboplatin was developed and clearly displays less nephro-, oto- and neurotoxicity. However, myelosuppression is more frequent, and as the drug is exlusively cleared through the kidney, adjustment of dose for creatinine clearance must be accomplished through use of various dosing nomograms.

Antitumor Antibiotics and Topoisomerase Poisons Antitumor antibiotics are substances produced by bacteria that in nature appear to provide chemical defense against other hostile microorganisms. As a class they bind to DNA directly and can frequently undergo electron transfer reactions to generate free radicals in close proximity to DNA, leading to DNA damage in the form of single strand breaks or cross-links. *Topoisomerase poisons* include natural products or semi-synthetic species derived ultimately from plants, and they modify enzymes that regulate the capacity of DNA to unwind to allow normal replication or transcription.

Doxorubicin is the most widely active and frequently used antineoplastic agent. It can intercalate into DNA, thereby altering DNA structure, replication, and topoisomerase function. It can also undergo redox cycling by accepting electrons into its quinone ring system. It causes predictable myelosuppression, alopecia, nausea, and mucositis. In addition, it causes acute cardiotoxicity in the form of atrial and ventricular dysrhythmias, but these are rarely of clinical significance. In contrast, cumulative doses >550 $mg/m_2$are associated with a 10% incidence of chronic cardiomyopathy. The incidence of cardiomyopathy appears to be related to schedule (peak serum concentration), with low dose, frequent treatment, or continuous infusions better tolerated than intermittent higher dose exposures. Radiation recall or interaction with radiation to cause local site complications is frequent. The drug is a powerful vesicant, with necrosis of tissue apparent 4 to 7 days after an extravasation; therefore it should be administered into a rapidly flowing intravenous line. The drug is metabolized by the liver, so doses must be reduced by 50 to 75% in the presence of liver dysfunction. Daunorubicin is closely related to doxorubicin and was actually introduced first into leukemia treament, where it remains part of curative regimens and has been shown preferable to doxorubicin owing to less mucositis and colonic damage. Idarubicin is an orally acting doxorubicin analogue, whose ultimate place in therapy is uncertain.

Bleomycin refers to a mixture of glycopeptides that have the unique feature of forming complexes with Fe^{2+}while bound to DNA. Oxidation gives rise to superoxide and hydroxyl radicals. The drug is of interest clinically as it causes little, if any, myelosuppression. The drug is cleared rapidly, but augmented skin and pulmonary

toxicity in the presence of renal failure has led to the recommendation that doses be reduced by 50 to 75% in the face of a creatinine clearance<25 mL/min. Bleomycin is not a vesicant, and can be administered intravenously, intramuscularly, or subcutaneously. Common side effects include fever and chills, facial flush, and Raynaud's syndrome. Hypertension can follow rapid intravenous administration, and the incidence of anaphylaxis with early preparations of the drug has led to the practice of administering a test dose of 0.5 to 1 unit before the rest of the dose. The most feared complication of bleomycin treatment is pulmonary fibrosis, which increases in incidence at >300 cumulative units administered and is at best minimally responsive to treatment (e.g., glucocorticoids). The earliest indicator of an adverse effect is a decline in the DL_{CO}, although cessation of drug immediately upon documentation of a decrease in DL_{CO} may not prevent further decline in pulmonary function. Bleomycin is inactivated by a bleomycin hydrolase, whose concentration is diminished in skin and lung. Because bleomycin-dependent electron transport is dependent on O_2, bleomycin toxicity may become apparent after exposure to transient very high inspired P_{O2}. Thus, during surgical procedures, patients with prior exposure to bleomycin should be maintained on the lowest inspired P_{O2}consistent with maintaining adequate tissue oxygenation.

D-Actinomycin intercalates into DNA and appears to have less, but not absent, capacity to undergo electron transfer reactions. It causes severe myelosuppression, nausea, alopecia, and mucositis. It is a notable vesicant. Mithramycin historically was used against testicular and other neoplasms; however, in addition to causing nausea, myelosuppression, and vesicant properties, it causes an "acute hemorrhagic syndrome" consisting of platelet function defects in association with indicators of disseminated intravascular coagulation. It is used in current practice to control hypercalcemia. In addition, renal and hepatic dysfunction may complicate its use.

Mitomycin C undergoes reduction of its quinone function to generate a bifunctional DNA alkylating agent. It is a broadly active antineoplastic agent with a number of unpredictable toxicities, including delayed bronchospasm 12 to 14 h after dosing and a chronic pulmonary fibrosis syndrome more frequent at doses of 50 to 60 mg/m_2. Cardiomyopathy has been described, particularly in the setting of prior radiation therapy. A hemolytic-uremic syndrome carries an ultimate mortality rate of 25 to 50% and is poorly treated by conventional component support and exchange transfusion. Mitomycin is a notable vesicant and causes substantial nausea and vomiting. It can be used for intravesical instillation for curative treatment of superficial transitional bladder carcinomas and, with radiation therapy, for curative treatment of anal carcinoma.

Mitoxantrone is a synthetic compound that was designed to recapitulate features of doxorubicin but with less cardiotoxicity. It is quantitatively less cardiotoxic (comparing the ratio of cardiotoxic to therapeutically effective doses), but its status in therapy is unclear as doses of 150 mg/$m_2$have produced evidence of 10% incidence of cardiotoxicity; it also causes alopecia.

Etoposide was synthetically derived from the plant product podophyllotoxin, and it binds directly to topoisomerase II and DNA in a reversible ternary complex. In that capacity, it stabilizes the covalent intermediate in the enzyme's action where the enzyme is covalently linked to DNA. This "alkali-labile" DNA bond was historically a first hint that an activity such as topoisomerase exists. The drug therefore causes a prominent G2 arrest,

reflecting the action of a DNA damage checkpoint. Prominent clinical effects include myelosuppression, nausea, and transient hypotension related to the speed of administration of the agent. Etoposide is a mild vesicant but is relatively free from other "large organ" toxicities. Teniposide is a structural relative with unique activity in childhood lymphoblastic leukemia. When given at high doses or very frequently, topoisomerase inhibitors may cause acute leukemia associated with chromosome 11q23 abnormalities in up to 1% of exposed patients.

Camptothecin was isolated from extracts of a Chinese tree and had notable antileukemia activity. Early clinical studies with the sodium salt of the hydrolyzed camptothecin lactone showed evidence of notable toxicity with little activity. Identification of topoisomerase I as the target of camptothecins and the need to preserve lactone structure allowed additional efforts to identify active members of this series. Topoisomerase I is responsible for unwinding the DNA strand by introducing single strand breaks and allowing rotation of one strand about the other. In S phase, topoisomerase I-induced breaks that are not promptly resealed lead to progress of the replication fork off the end of a DNA strand. The DNA damage is a potent signal for induction of apoptosis. Camptothecins promote the stabilization of the DNA linked to the enzyme in a so-called cleavable complex, analogous to the action of etoposide with topisomerase II. Topotecan is a camptothecin derivative approved for use in ovarian tumors. Toxicity is limited to myelosuppression and mucositis. CPT-11, or irinotecan, is a camptothecin with evidence of activity in colon carcinoma. In addition to myelosuppression, it causes a secretory diarrhea, which can be treated effectively with loperamide or octreotide.

Indirect Effectors of DNA Function: Antimetabolites A broad definition of antimetabolites would include compounds with structural similarity to precursors of purines or pyrimidines or that interfere with purine or pyrimidine synthesis. Antimetabolites can cause DNA damage indirectly, through misincorporation into DNA, abnormal timing or progression through DNA synthesis, or altered function of pyrimidine and purine biosynthetic enzymes. They tend to convey greatest toxicity to cells in S phase, and the degree of toxicity increases with duration of exposure. Common toxic manifestations include stomatitis, diarrhea, and myelosuppression. Second malignancies are not associated with their use.

Methotrexate inhibits dihydrofolate reductase, which regenerates reduced folates from the oxidized folates produced when thymidine monophosphate is formed from deoxyuridine monophosphate. Without reduced folates, cells die a "thymineless" death. N_5tetrahydrofolate or N_5formyltetrahydrofolate (leucovorin) can bypass this block and rescue cells from methotrexate, which is maintained in cells by polyglutamylation. The drug and other reduced folates are transported into cells by the folate carrier, and high concentrations of drug can bypass this carrier and allow diffusion of drug directly into cells. These properties have suggested the design of "high-dose" methotrexate regimens with leucovorin rescue of normal marrow and mucosa, part of curative approaches to osteosarcoma in the adjuvant setting, and hematopoietic neoplasms of children and adults. Methotrexate is cleared by the kidney by both glomerular filtration and tubular secretion, and toxicity is augmented by renal dysfunction and drugs such as salicylates, probenecid, and nonsteroidal anti-inflammatory agents that undergo tubular secretion. With normal renal function, 15 mg/$m_2$leucovorin will rescue $10^{-8}$to 10^{-6}*M*

methotrexate in three to four doses. However, with decreased creatinine clearance, doses of 50 to100 mg/$m^2$are continued until methotrexate levels are $<5\acute{}10^{-8}M$. In addition to bone marrow suppression and mucosal irritation, methotrexate can cause renal failure itself at high doses owing to crystallization in renal tubules; therefore high-dose regimens require alkalinization of urine with increased flow by hydration. Methotrexate can be sequestered in third space collections and leech back into the general circulation, causing prolonged myelosuppression. Less frequent adverse effects include reversible increases in transaminases and a hypersensitivity-like pulmonary syndrome. Chronic low-dose methotrexate can cause hepatic fibrosis. When administered to the intrathecal space, methotrexate can cause chemical arachnoiditis andCNSdysfunction. Trimetrexate is a methotrexate derivative that is not polyglutamylated and does not use the reduced folate carrier.

5-Fluorouracil (5FU) represents an early example of "rational" drug design in that it originated from the observation that tumor cells incorporate uracil more efficiently into DNA than normal cells, especially gut. 5FU is metabolized in cells to 5¢FdUMP, which inhibits thymidylate synthetase (TS). In addition, misincorporation can lead to single strand breaks, and RNA can aberrantly incorporate FUMP. 5FU is metabolized by dihydropyrimidine dehydrogenase, and deficiency of this enzyme can lead to excessive toxicity from 5FU. Oral bioavailability varies unreliably. Intravenous administration leads to bone marrow suppression after short infusions but to more evidence of stomatitis after prolonged infusions. Leucovorin augments the toxicity and activity of 5FU by promoting formation of the ternary covalent complex of 5FU, the reduced folate, and TS. Less frequent toxicities includeCNSdysfunction, with prominent cerebellar signs, and endothelial toxicity manifested by thrombosis, including pulmonary embolus and myocardial infarction.

Cytosine arabinoside (ara-C) is incorporated into DNA after formation of ara-CTP, resulting in S phase-related toxicity. Continuous infusion schedules allow maximal efficiency of effect, with uptake maximal at 5 to 7 u*M*. Ara-C can be administered intrathecally. Adverse effects include nausea, diarrhea, stomatitis, chemical conjunctivitis, and cerebellar ataxia. Gemcitabine is a cytosine derivative that is similar to ara-C in that it is incorporated into DNA after anabolism to the triphosphate, rendering DNA susceptible to breakage and repair synthesis, which differs from that in ara-C in that lesions including the analogue are very inefficiently removed. In contrast to ara-C, gemcitabine appears to have useful activity in a variety of solid tumors, with limited nonmyelosuppressive toxicities. 6-Thioguanine and 6-mercaptopurine (6MP) are used in the treatment of acute lymphoid leukemia. Although administered orally, they display very variable bioavailability. 6MP is metabolized by xanthine oxidase and therefore requires dose reduction when used with allopurinol.

Fludarabine phosphate is a prodrug of F-ara-A, which in turn was designed to diminish the susceptibility of ara-A to adenosine deaminase. Ara-A is incorporated into DNA and can cause delayed cytotoxicity even in cells with low growth fraction, including chronic lymphocytic leukemia and follicular B cell lymphoma.CNSdysfunction and T cell depletion leading to opportunistic infections can occur in addition to myelosuppression. 2-Chlorodeoxyadenosine is a similar compound with activity in hairy cell leukemia. 2-Deoxycoformycin inhibits adenosine deaminase, with resulting increase in dATP levels. This causes inhibition of ribonucleotide reductase as well as augmented

susceptibility to apoptosis, particularly in T cells. Renal failure and CNS dysfunction are notable toxicities in addition to immunosuppression.

Hydroxyurea inhibits ribonucleotide reductase, resulting in S phase block. It is orally bioavailable and the drug of choice for the acute management of myeloproliferative states. Asparaginase is not classically considered an antimetabolite as it causes breakdown of extracellular asparagine required for protein synthesis in certain leukemic cells. However, it effectively stops DNA synthesis by preventing the requisite concurrent protein synthesis, and therefore it has a similar functional outcome as the classic antimetabolites. As asparaginase is a foreign protein, hypersensitivity reactions are common, as are effects on organs such as pancreas and liver that require continuing protein synthesis. This results in decreased insulin secretion with hyperglycemia, with or without hyperamylasemia and clotting function abnormalities. The latter may be associated withCNS and dural vein thrombosis.

Mitotic Spindle Inhibitors Microtubules are cellular structures that form the mitotic spindle and in interphase cells are responsible for the cellular "scaffolding" along which various motile and secretory processes occur. Microtubules are composed of repeating noncovalent multimers of a hetrodimer ofa and b subunits of the protein tubulin. Vincristine binds to the tubulin dimer with the result that microtubules are disaggregated. This results in the block of growing cells in M phase; however, toxic effects in G1 and S phase are also evident. The drug is bound to blood-formed elements, leading to its occasional use as vinca-loaded platelets to treat autoimmune thrombocytopenia. The drug is metabolized by the liver, and dose adjustment in the presence of hepatic dysfunction is required. It is a powerful vesicant, and infiltration can be treated by local heat and infiltration with hyaluronidase. The drug is lethal if inadvertantly administered by the intrathecal route. At clinically used intravenous doses, neurotoxicity in the form of glove-and-stocking neuropathy is frequent. Children tolerate 2 mg/m_2, but adult doses may be capped at 2 mg total to lower the incidence of disabling chronic neuropathy; whether this compromises needed dose intensity in curative regimens is uncertain. Acute neuropathic effects include jaw pain, paralytic ileus, urinary retention, and the syndrome of inappropriate antidiuretic hormone secretion. Myelosuppression is not seen. Vinblastine is similar to vincristine, except that it tends to be more myelotoxic, with more frequent thrombocytopenia and also mucositis and stomatitis. Vinorelbine is a recently introduced vinca alkaloid that appears to have differences in resistance patterns in comparison to vincristine and vinblastine; it may be administered orally.

The taxanes include paclitaxel and docetaxel. These agents differ from the vinca alkaloids in that the taxanes stabilize microtubules against depolymerization. The "stabilized" microtubules function abnormally and are not able to undergo the normal dynamic changes of microtubule function necessary for cell cycle completion. Taxanes are among the most broadly active antineoplastic agents for use in solid tumors, with evidence of activity in ovarian cancer, breast cancer, Kaposi's sarcoma, and lung tumors. They are administered intravenously, and paclitaxel requires use of a cremophore-containing vehicle that can cause hypersensitiviy reactions. Premedication with regimens including dexamethasone (20 mg orally or intravenously 12 and 6 h before treatment), diphenhydramine (50 mg), and cimetidine (300 mg) both 30 min before treatment decreases but does not eliminate the risk of hypersensitivity reactions to the paclitaxel vehicle. Docetaxel uses a polysorbate 80 formulation, which can cause

fluid retention in addition to hypersensitivity reactions, and dexamethasone premedication with or without antihistamines is frequently used. Paclitaxel causes hypersensitivity reactions, myelosuppression, neurotoxicity in the form of glove-and-stocking numbness, and paresthesia. Cardiac rhythm disturbances were observed in phase I and II trials, most commonly asymptomatic bradycardia but, much more rarely, varying degrees of heart block. These have not emerged as clinically significant in the majority of patients. Infrequently occurring evidence of myocardial ischemia during paclitaxel administration cannot yet be clearly related to the drug. Docetaxel causes comparable degrees of myelosuppression and neuropathy. Hypersensitivity reactions, including bronchospasm, dyspnea, and hypotension, are less frequent but occur to some degree in up to 25% of patients. Fluid retention appears to result from a vascular leak syndrome that can aggravate preexisting effusions. Rash can complicate docetaxel administration, appearing prominently as a pruritic maculopapular rash affecting the forearms, but it has also been associated with fingernail ridging, breakdown, and skin discoloration. Stomatitis appears to be somewhat more frequent than with paclitaxel.

Estramustine was originally synthesized as a mustard derivative that might be useful in neoplasms that possessed estrogen receptor sites. However, no evidence of interaction with DNA was observed. Surprisingly, the drug caused metaphase arrest, and subsequent study revealed that it binds to microtubule-associated proteins, resulting in abnormal microtubule function. Estramustine binds to estramustine-binding proteins (EMBP), which have particular presence in prostate tumor tissue. The drug is approved for treatment of metastatic prostate cancer as an oral formulation. Gastrointestinal and cardiovascular adverse effects related to the estrogen moiety occur in up to 10% of patients, including worsened heart failure and thromboembolic phenomena. Gynecomastia and nipple tenderness can also occur.

Hormonal Agents The family of steroid hormone receptor-related molecules have emerged as prominent targets for "small molecules" useful in cancer treatment. When bound to their cognate ligands, these receptors can alter gene transcription and, in certain tissues, induce apoptosis. The pharmacologic effect is a mirror or parody of the normal effects of the agent acting in nontransformed tissue, although the effects on tumors are mediated by indirect effects in some cases.

Glucocorticoids are generally given in "pulsed" high-dose exposure in leukemias and lymphomas, where they induce apoptosis in tumor cells. Cushing's syndrome or inadvertent adrenal suppression on withdrawal from high-dose glucocorticoids can be significant complications, along with infections common in immunosuppressed patients, in particular *Pneumocystis* pneumonia, which classically appears a few days after completing a course of high-dose steroids. Tamoxifen is a partial estrogen receptor antagonist; it has a tenfold greater degree of antitumor activity in breast cancer patients whose tumors express estrogen receptors than in those who have low or no levels of expression. Side effects include a somewhat increased risk of estrogen-related cardiovascular complications, such as thromboembolic phenomena, and a small increased incidence of endometrial carcinoma, which appears after chronic use. Progestational agents including medroxyprogesterone acetate, androgens including halotestin, and, paradoxically, estrogens have approximately the same degeree of activity in primary hormonal treatment of breast cancers that have elevated levels of

estrogen receptors. Estrogen is not used often owing to prominent cardiovascular and uterotropic activity.

Prostate cancer is classically treated by diethylstilbesterol (DES) acting as an estrogen at the level of the hypothalamus to downregulate hypothalamic luteinizing hormone (LH) production, resulting in decreased elaboration of testosterone by the testicle. For this reason, orchiectomy is equally as effective as moderate-dose DES, inducing responses in ~80% of previously untreated patients with prostate cancer but without the prominent cardiovascular side effects of DES, including thrombosis and exacerbation of coronary artery disease. In the event that orchiectomy is not accepted by the patient, testicular androgen suppression can also be effected by luteinizing hormone-releasing hormone (LHRH) agonists such as leuprolide and goserelin. These agents cause tonic stimulation of the LHRH receptor, with the loss of its normal pulsatile activation resulting in its desensitization and decreased output of LH by the anterior pituitary. Therefore, as primary hormonal manipulation in prostate cancer one can choose orchiectomy or leuprolide, not both. The addition of actual antagonists of androgens acting at the androgen receptor, including flutamide or bicalutamide, is of uncertain additional benefit in extending overall response duration, but it clearly prevents the activation of androgen receptors by adrenal androgens, and the combined use of orchiectomy or leuprolide plus flutamide is referred to as "total androgen blockade."

Interestingly, tumors that respond to a primary hormonal manipulation may frequently respond to second and third hormonal manipulations. Thus, breast tumors that had previously responded to tamoxifen have, on relapse, notable response rates to withdrawal of tamoxifen itself or to subsequent addition of a progestin. Likewise, initial treatment of prostate cancers with leuprolide plus flutamide may be followed after disease progression by response to withrawal of flutamide. These responses may result from the removal of antagonists from mutant steroid hormone receptors that have come to depend on the presence of the antagonist as a growth-promoting influence.

Additional strategies to treat refractory breast and prostate cancers that possess steroid hormone receptors may also address adrenal capacity to produce androgens and estrogens, even after orchiectomy or oophorectomy, respectively. Thus, aminoglutethimide or ketoconazole can be used to block adrenal synthesis by interfering with the enzymes of steroid hormone metabolism. Administration of these agents requires concomitant hydrocortisone replacement and additional glucocorticoid doses administered in the event of physiologic stress. Steroid hormone-inducing "aromatase" activity may be present in tumor tissue, and second- or third-line approaches to inhibition of aromatase activity may also be effected by such agents as anastrazole and letrozole.

Humoral mechanisms can also result in complications of an underlying malignancy. Adrenocortical carcinomas can cause Cushing's syndrome as well as syndromes of androgen or estrogen excess. Mitotane can counteract these by decreasing synthesis of steroid hormones. Islet cell neoplasms can cause debilitating diarrhea, treated with the somatostatin analogue octreotide. Prolactin-secreting tumors can be effectively managed by the dopaminergic agononist bromocriptine.

An additional strategy related conceptually to the use of steroid hormones is the use of

retinoids, including tretinoin, the all-*trans*-isomer of retinoic acid, or isotretinoin, the 13-*cis* isomer of retinoic acid, to cause "differentiation" by acting on the retinoid receptor, which is in the steroid hormone receptor family. In particular, tretinoin is part of curative regimens for acute promyelocytic leukemia (PML) and appears to act by causing accelerated degradation of the fusion protein created by the t(15;17) translocation fusing the retinoic acid receptor a and the PML transcription factor. Acute side effects related to differentiation of promyelocytes to mature granulocytes may result in pulmonary symptoms related to granulocyte sequestration in the pulmonary vasculature; these are treated by respiratory support and glucocorticoids. Squamous neoplasms, including those of the skin and cervix, also appear to be uniquely responsive in certain cases to retinoids.

ACUTE COMPLICATIONS OF CANCER CHEMOTHERAPY

Myelosuppression The common cytotoxic chemotherapeutic agents almost invariably affect bone marrow function. Titration of this effect determines in many cases the MTD of the agent on a given schedule. The normal kinetics of blood cell turnover influence the sequence and sensitivity of each of the formed elements. Polymorphonuclear leukocytes (PMNs; $T_{1/2}$= 6 to 8 h), platelets ($T_{1/2}$= 5 to 7 days), and red blood cells (RBC; $T_{1/2}$~ 120 days) have most, less, and least susceptibility to usually administered cytotoxic agents, respectively. The *nadir count* of each cell type in response to classes of agents is characteristic. Maximal neutropenia occurs 6 to 14 days after conventional doses of anthracyclines, antifolates, and antimetabolites. Alkylating agents differ in timing of cytopenias. Nitrosoureas, DTIC, and procarbazine can display delayed marrow toxicity, first appearing 6 weeks after dosing.

Complications of myelosuppression result from the predictable sequelae of the missing cells' function. *Febrile neutropenia* refers to the clinical presentation of fever (one temperature³38.5°C or three readings ³38°C but£38.5°C per 24 h) in a cytopenic patient with an uncontrolled neoplasm involving the bone marrow or, more usually, in a patient undergoing treatment with cytotoxic agents. Mortality from uncontrolled infection varies inversely with thePMNcount. If the nadir PMN count is >1000/uL, there is little risk; if <500/uL, risk of death is markedly increased. Management of febrile neutropenia has conventionally included empirical coverage with antibiotics for the duration of neutropenia (Chap. 85). Selection of antibiotics is governed by the expected association of infections with certain underlying neoplasms; careful physical examination (with scrutiny of catheter sites, dentition, mucosal surfaces, and perirectal and genital orifices by gentle palpation); chest x-ray; and Gram stain and culture of blood, urine and sputum (if any) to define a putative site of infection. In the absence of any originating site, a broadly acting b-lactam with anti-*Pseudomonas* activity, such as ceftazidime, is begun empirically. The addition of vancomycin to cover potential cutaneous sites of origin (until these are ruled out or shown to originate from methicillin-sensitive organisms) or metronidazole or imipenem for abdominal or other sites favoring anaerobes reflects modifications tailored to individual patient presentations. The coexistence of pulmonary compromise raises a distinct set of potential pathogens, including *Legionella*, *Pneumocystis*, and fungal agents, that may require further diagnostic evaluations such as bronchoscopy with bronchoalveolar lavage. Febrile neutropenic patients can be stratified broadly into two prognostic groups. The first, with expected short duration of neutropenia and no evidence of hypotension or abdominal or other localizing symptoms,

may be expected to do well even with less complex, oral regimens, e.g., ciprofloxacin plus amoxicillin/clavulinic acid. Detailed evaluation of such simple oral programs and intravenous regimens is ongoing. A less favorable prognostic group are patients with expected prolonged neutropenia, evidence of sepsis, and end-organ compromise, particularly pneumonia. These patients clearly require tailoring of their antibiotic regimen to their underlying presentation, with frequent empirical addition of antifungal agents if fever persists for 7 days without identification of an adequately treated organism or site.

Transfusion of granulocytes has no role in the management of febrile neutropenia, owing to their exceedingly short half-life, mechanical fragility, and clinical syndromes of pulmonary compromise with leukostasis after their use. Instead, *colony stimulating factors* (CSFs) are used to augment bone marrow production ofPMNs. These include "early-acting" factors such asIL-1, IL-3, and stem cell factor, which act on multiple lineages, and "late-acting" lineage-specific factors such as G-CSF (granulocyte colony stimulating factor) or GM-CSF (granulocyte-macrophage colony stimulating factor), erythropoietin, thrombopoietin, IL-6, and IL-11. CSFs are overused in oncology practice. The settings in which their use has been proven effective are limited. G-CSF, GM-CSF, erythropoietin, and IL-11 are currently approved for use. The American Society of Clinical Oncology has developed practice guidelines for the use of G-CSF and GM-CSF (Table 84-3). Primary administration (i.e., shortly after completing chemotherapy to reduce the nadir) of G-CSF to patients receiving cytotoxic regimens associated with a 40% incidence of febrile neutropenia has reduced the incidence of febrile neutropenia in several studies by about 50%. Most patients, however, receive regimens that do not have such a high risk of expected febrile neutropenia, and therefore most patients initially should not receive G-CSF or GM-CSF. Special circumstances such as a documented history of febrile neutropenia with the regimen in a particular patient; extensive compromise of marrow by prior radiation or chemotherapy; or active, open wounds or deep-seated infection may support primary treatment with G-CSF or GM-CSF. Administration of G-CSF or GM-CSF to afebrile neutropenic patients or to patients with "low-risk" febrile neutropenia (secondary administration, use after neutropenia has developed) as defined above is not recommended, although administration to "high-risk" patients with febrile neutropenia and evidence of organ compromise is reasonable. G-CSF or GM-CSF is conventionally started 24 to 72 h after completion of chemotherapy and continued until a PMN count of 10,000/uL is achieved. Also, patients with myeloid leukemias undergoing induction therapy may have a slight reduction in the duration of neutropenia if G-CSF (not GM-CSF) is commenced after completion of therapy and may be of particular value in elderly patients, but the influence on long-term outcome has not been defined. GM-CSF probably has a more restricted utility than G-CSF, with its use currently limited to patients after autologous bone marrow transplants, although proper "head-to-head" comparisons with G-CSF have not been conducted in most instances. GM-CSF may be associated with more systemic side effects.

Dangerous degrees of thrombocytopenia do not frequently complicate the management of patients with solid tumors receiving cytotoxic chemotherapy (carboplatin-containing regimens are frequently involved), but they are frequent in patients with certain hematologic neoplasms where marrow is infiltrated with tumor. Severe bleeding related to thrombocytopenia occurs with increased frequency at platelet counts<20,000/uL and is very prevalent at counts<5000/uL. Prophylactic transfusions to keep platelets

>20,000/uL are warranted in patients with leukemia (the threshold for transfusion is 10,000/uL in patients with solid tumors and no other bleeding diathesis). Careful review of medication lists to prevent exposure to nonsteroidal anti-inflammatory agents and maintenance of clotting factor levels adequate to support near-normal prothrombin and partial thromboplastin time tests are of import in minimizing the risk of bleeding in the thrombocytopenic patient. Certain cytokines in clinical investigation have shown ability to increase platelets (e.g.,IL-6, IL-1, thrombopoietin), but clinical benefit and safety are not yet proven. IL-11 (oprelvekin) is approved for use in the setting of expected thrombocytopenia, but its effects on platelet counts are small and it is associated with side effects such as headache, fever, malaise, syncope, cardiac arrhythmias, and fluid retention.

Anemia associated with chemotherapy can be managed by transfusion of packedRBCs. Transfusion is not undertaken until the hemoglobin falls to<80 g/L (8 g/dL), or if compromise of end-organ function occurs or an underlying condition (e.g., coronary artery disease) calls for maintenance of hemoglobin >90 g/L (9 g/dL). Patients who are to receive therapy for >2 months on a "stable" regimen and who are likely to require continuing transfusions are also candidates for erythropoietin to maintain hemoglobin of 90 to 100 g/L (9 to 10 g/dL). In the setting of adequate iron stores and serum erythropoietin levels<100 ng/mL, erythropoietin, 150 U three times a week, can produce a slow increase in hemoglobin over about 2 months of administration. Quality of life is better at higher hemoglobin concentrations, but expense is a concern with erythropoietin use.

Nausea and Vomiting The most common side effect of chemotherapy administration is nausea, with or without vomiting. Antineoplastic agents vary in their capacity to cause nausea and vomiting. Nitrogen mustard, nitrosoureas, streptozotocin, DTIC, cisplatin, and actinomycin are highly emetogenic and produce vomiting in virtually all patients. Doxorubicin, daunorubicin, and conventional-dose cyclophosphamide are moderately emetogenic. Antimetabolites are dose- and schedule-dependent, with single doses of methotrexate and fluorouracil producing at worst anorexia; while 5-day regimens of 5-fluorouracil and high-dose methotrexate produce nausea in ~50% of patients. Other agents such as chlorambucil, melphalan, and busulfan in conventional doses produce little tendency to emesis.

Emesis is a reflex caused by stimulation of the vomiting center in the medulla. Input to the vomiting center comes from the chemoreceptor trigger zone (CTZ) and afferents from the peripheral gastrointestinal tract, cerebral cortex, and heart. In addition, a conditioned reflex may contribute to anticipatory nausea arising after repeated cycles of chemotherapy. Accordingly, antiemesis agents differ in their locus of action. Combining agents from different classes or the sequential use of different classes of agent is the cornerstone of successful management of chemotherapy-induced nausea and vomiting. Of great importance are the prophylactic administration of agents and the use of psychological techniques including the maintenance of a supportive milieu, counseling, and relaxation to augment the action of antiemetic agents.

Antidopaminergic phenothiazines act directly at theCTZ, and include prochlorperazine (Compazine), 10 mg intramuscularly or intravenously, 10 to 25 mg orally, or 25 mg per rectum every 4 to 6 h for up to four doses; and thiethylperazine (Torecan), 10 mg by all

the above routes every 6 h. Haloperidol (Haldol) is a butyrophenone dopamine antagonist given at 0.5 to 1.0 mg intramuscularly or orally every 8 h. Antihistamines such as diphenhydramine (Benadryl) have little intrinsic antiemetic capacity but are frequently given to prevent or treat dystonic reactions that can complicate use of the antidopaminergic agents. Lorazepam (Ativan) is a short-acting benzodiazepine that provides an anxiolytic effect to augment the effectiveness of a variety of agents when used at 1 to 2 mg intramuscularly, intravenously, or orally every 4 to 6 h. Dexamethasone (Decadron) likewise augments the action of a variety of agents when used at 4 to 40 mg intravenously or orally, given before treatment and repeated up to 10 mg orally every 6 h four times. Metoclopramide (Reglan) acts on peripheral dopamine receptors to augment gastric emptying and is used in high doses for highly emetogenic regimens (1 to 2 mg/kg intravenously 30 min before chemotherapy and every 2 h for up to three additional doses as needed); intravenous doses of 10 to 20 mg every 4 to 6 h as needed or 50 mg orally 4 h before and 8 and 12 h after chemotherapy are used for moderately emetogenic regimens. Serotonin antagonists are useful in moderately to severely emetogenic regimens; ondansetron (Zofran) is given as 0.15 mg/kg intravenously for three doses just before and at 4 and 8 h after chemotherapy, and granisetron (Kytril) is given as a single dose of 0.01 mg/kg just before chemotherapy.d-9-Tetrahydrocannabinol (Marinol) is a rather weak antiemetic compared to other available agents, but it may be useful for persisting nausea and is used orally at 10 mg every 3 to 4 h as needed.

Alopecia Chemotherapeutic agents vary widely in causing alopecia, with anthracyclines, alkylating agents, and topoisomerase inhibitors reliably causing near total alopecia when given at therapeutic doses. Antimetabolites are more variably associated with alopecia. Psychologic support and the use of cosmetic resources are to be encouraged, and "chemo caps" that reduce scalp temperature to decrease the degree of alopecia should be discouraged.

Gonadal Dysfunction and Pregnancy Cessation of ovulation and azoospermia reliably result from alkylating agent- and topoisomerase poison-containing regimens. The duration of these effects varies with age and sex. Males treated for Hodgkin's disease with mechlorethamine- and procarbazine-containing regimens are effectively sterile, while fertility usually returns after regimens including cisplatin, vinblastine or etoposide, and bleomycin for testicular cancer. Sperm banking before treatment may be considered to support patients likely to be sterilized by treatment. Females experience amenorrhea with anovulation after alkylating agent therapy but are likely to recover normal menses if treatment is completed before age 30 and unlikely to recover menses after age 35. Even those who regain menses usually experience premature menopause. As the magnitude and extent of decreased fertility can be difficult to predict, patients should be counseled to maintain effective contraception, preferably by barrier means, during and after therapy. Resumption of efforts to conceive should be considered in the context of the likely prognosis of the patient. Hormone-replacement therapy should be undertaken in women who do not have a hormonally responsive tumor.

Chemotherapy agents have variable effects on the success of pregnancy (Chap. 7). All agents tend to have increased risk of adverse outcomes when administered during the first trimester, and strategies to delay chemotherapy if possible until after this milestone should be considered if the pregnancy is to continue to term. Patients in their second or

third trimester can be treated with most regimens for the common neoplasms afflicting women in their child-bearing years with the exception of antimetabolites, particularly antifolates, which have notable teratogenic or fetotoxic effects throughout pregnancy. The need for anticancer chemotherapy per se is infrequently a clear basis to recommend termination of a concurrent pregnancy, although each treatment strategy in this circumstance must be tailored to the individual needs of the patient.*Chronic effects of cancer treatment are reviewed in Chap. 103.*

BIOLOGIC THERAPY

No postulates resembling principles have emerged from efforts to develop biologic approaches to cancer treatment. The goal of biologic therapy is to manipulate the host-tumor interaction in favor of the host. Theoretically, biologic approaches should reflect a bell-shaped dose-response curve where the maximum biologic effect is less than the MTD. Empirical trial and error has led to the discovery that a number of biologic treatment approaches may produce antitumor effects, but nearly all of them are most active at their MTD.

IMMUNE MEDIATORS OF ANTITUMOR EFFECTS

The very existence of a cancer in a person is testimony to the failure of the immune system to deal effectively with the cancer. Tumors have a variety of means of avoiding the immune system: (1) they are often only subtly different from their normal counterparts; (2) they are capable of downregulating their major histocompatibility complex antigens, effectively masking them from recognition by T cells; (3) they are inefficient at presenting antigens to the immune system; (4) they can cloak themselves in a protective shell of fibrin to minimize contact with surveillance mechanisms; and (5) they can produce a range of soluble molecules, including potential immune targets, that can distract the immune system from recognizing the tumor cell. Some of the cell products initially polarize the immune response away from cellular immunity (shifting from Th1 to Th2 responses, Chap. 305) and ultimately lead to defects in T cells that prevent their activation and cytotoxic activity. Cancer treatment further suppresses host immunity. A variety of strategies are being tested to overcome these barriers.

Cell-Mediated Immunity The strongest evidence that the immune system can exert clinically meaningful antitumor effects comes from allogeneic bone marrow transplantation. Adoptively transferred T cells from the donor expand in the tumor-bearing host, recognize the tumor as being foreign, and mediate impressive antitumor effects (graft-versus-tumor effects). Three types of experimental interventions are being developed to take advantage of the ability of T cells to kill tumor cells.

1. Allogeneic T cells are being transferred to cancer-bearing hosts in three major settings: in the form of allogeneic bone marrow transplantation, as pure lymphocyte transfusions following bone marrow recovery after allogeneic bone marrow transplantation, and as pure lymphocyte transfusions following immunosuppressive (but not myeloablative) therapy (so-called minitransplants). In each of these settings, the effector cells are donor T cells that recognize the tumor as being foreign, probably through minor histocompatibility differences. The main risk of such therapy is the development of graft-versus-host disease because of the minimal difference between

the cancer and the normal host cells. This approach has been highly effective in hematologic cancers.

2. Autologous T cells are being removed from the tumor-bearing host, manipulated in several ways in vitro, and given back to the patient. The two major classes of autologous T cell manipulation are: (1) to develop tumor antigen-specific T cells and expand them to large numbers over many weeks ex vivo before administration, and (2) to activate the cells with polyclonal stimulators such as anti-CD3 and anti-CD28 after a short period ex vivo and try to expand them in the host after adoptive transfer with stimulation byIL-2, for example. Short periods removed from the patient permit the cells to overcome the tumor-induced T cell defects, and such cells traffic and home to sites of disease better than cells that have been in culture for many weeks. Individual centers have successful experiences with one or the other approach but not both, and whether one is superior to the other is not known.

3. Tumor vaccines are aimed at boosting T cell immunity. The finding that mutant oncogenes that are expressed only intracellularly can be recognized as targets of T cell killing greatly expanded the possibilities for tumor vaccine development. No longer is it difficult to find something different about tumor cells from normal cells. However, major difficulties remain in getting the tumor-specific peptides presented in a fashion to prime the T cells. Tumors themselves are very poor at presenting their own antigens to T cells at the first antigen exposure (*priming*). Priming is best accomplished by professional antigen-presenting cells (dendritic cells). Thus, a number of experimental strategies are aimed at priming host T cells against tumor-associated peptides. Vaccine adjuvants such asGM-CSFappear capable of attracting antigen-presenting cells to a skin site containing a tumor antigen. Such an approach has been documented to eradicate microscopic residual disease in follicular lymphoma and give rise to tumor-specific T cells. Purified antigen-presenting cells can be pulsed with tumor, its membranes, or particular tumor antigens and delivered as a vaccine. Tumor cells can be transfected with genes that attract antigen-presenting cells. Other ideas are also being tested. In a variation on the theme of adoptive transfer, the tumor vaccine may be given to the normal bone marrow and lymphoid cell donor of an allogeneic transplant so that the donor immune system has more cells capable of recognizing the tumor specifically. Vaccines against viral cancers (papillomavirus in cervical cancer), lymphomas, and melanomas have had modest clinical success.

Antibodies In general, antibodies are not very effective at killing cancer cells. Because the tumor seems to influence the host toward making antibodies rather than generating cellular immunity, it is inferred that antibodies are easier to defend against. Many patients can be shown to have serum antibodies directed at their tumors, but these do not appear to influence disease progression. However, the ability to grow very large quantities of high-affinity antibody directed at a tumor by the hybridoma technique has led to the application of antibodies to the treatment of cancer. The first study of a monoclonal antibody in cancer was published in 1980 and demonstrated many hurdles that needed to be overcome to make the approach successful. It seemed best to attack a determinant that was not shed or modulated by the tumor. A target that was involved in an important function for the tumor cells might be superior to a physiologically irrelevant target. Murine antibodies were not very effective because they did not mediate human effector mechanisms well and the host nearly always made antibodies against

the therapeutic antibody that prevented it from finding the target.

The lessons were learned; humanized antibodies against the CD20 molecule expressed on B cell lymphomas (rituximab) and against the HER-2/neu receptor overexpressed on epithelial cancers, especially breast cancer (herceptin), have become reliable tools in the oncologists armamentarium. Each used alone can cause tumor regression (rituximab> herceptin), and both appear to potentiate the effects of combination chemotherapy given just after antibody administration. It is likely that other antibodies against other important tumor targets will be available soon. Conjugation to drugs, toxins, isotopes, photodynamic agents, and other killing moieties may also be effective. Radioconjugates are the closest to approval. Other conjugates are associated with problems that have not yet been solved (e.g., antigenicity, instability, poor tumor penetration).

Cytokines There are>70 separate proteins and glycoproteins with biologic effects in humans:IFN-a, -b, -g;IL-1 through -18 (so far); the tumor necrosis factor (TFN) family [including lymphotoxin, TFN-related apoptosis-inducing ligand (TRAIL), CD40 ligand, and others]; and the chemokine family. Only a fraction of these has been tested against cancer; only IFN-a and IL-2 are in routine clinical use.

About 20 different genes encodeIFN-a, and their biologic effects are indistinguishable. Interferon induces the expression of many genes, inhibits protein synthesis, and exerts a number of different effects on diverse cellular processes. Its antitumor effects appear to be antagonized in vitro by thymidine, suggesting that de novo thymidylate synthesis is also affected. The two recombinant forms that are commercially available are IFN-a2a and -a2b. In general, interferon antitumor effects are dose-related, and IFN is most effective at itsMTD. Interferon is not curative for any tumor but can induce partial responses in follicular lymphoma, hairy cell leukemia, chronic myeloid leukemia, melanoma, and Kaposi's sarcoma. It has been used in the adjuvant setting in stage II melanoma, multiple myeloma, and follicular lymphoma. Its effects on survival are controversial. It produces fever, fatigue, a flulike syndrome, malaise, myelosuppression, and depression and can induce clinically significant autoimmune disease.

IL-2 must exert its antitumor effects indirectly through augmentation of immune function. Its biologic activity is to promote the growth and activity of T cells and natural killer (NK) cells. High doses of IL-2 can produce tumor regressions in ~20% of patients with metastatic melanoma and renal cell cancer. About 5% of patients may experience complete remissions that are durable, unlike any other treatment for these tumors. IL-2 is associated with myriad clinical side effects: intravascular volume depletion, capillary leak syndrome, adult respiratory distress syndrome, hypotension, fever, chills, skin rash, and impaired renal and liver function. Patients may require blood pressure support and intensive care to manage the toxicity. However, once the agent is stopped, most of the toxicities reverse completely within 3 to 6 days.

(Bibliography omitted in Palm version)

Back to Table of Contents

85. INFECTIONS IN PATIENTS WITH CANCER - *Robert Finberg*

Infections are a common cause of death and an even more common cause of morbidity in patients with a wide variety of neoplasms. Autopsy studies show that most deaths from acute leukemia and half of deaths from lymphoma are caused directly by infection. With more intensive chemotherapy, patients with solid tumors have become more likely to die of infection rather than their underlying disease.

A physical predisposition to infection (Table 85-1) can be a result of the neoplasm's production of a break in the skin; for example, a squamous cell carcinoma may cause local invasion of the epidermis, which allows bacteria to gain access to the subcutaneous tissue and permits the development of cellulitis. The artificial closing of a normally patent orifice can also predispose to infection: Obstruction of a ureter by a tumor can cause urinary tract infection, and obstruction of the bile duct can cause cholangitis. Part of the host's normal defense against infection depends on the continuous emptying of a viscus; without emptying, a few bacteria present as a result of bacteremia or local transit can multiply and cause disease.

A similar problem can affect patients whose lymph node integrity has been disrupted by radical surgery, particularly patients who have had radical node dissections. A common clinical problem following radical mastectomy is the development of cellulitis (usually caused by streptococci or staphylococci) because of lymphedema and/or inadequate lymph drainage. In most cases, this problem can be addressed by local measures designed to prevent fluid accumulation and breaks in the skin, but antibiotic prophylaxis has been necessary in refractory cases.

A life-threatening problem common to many cancer patients is the loss of the reticuloendothelial capacity to clear microorganisms after splenectomy. Splenectomy is common in patients with Hodgkin's disease and in the management of hairy cell leukemia, chronic lymphocytic leukemia (CLL), and refractory idiopathic thrombocytopenic purpura. Even after curative therapy for the underlying disease, the lack of a spleen predisposes such patients to rapidly fatal infections. The loss of the spleen through trauma similarly predisposes the normal host to overwhelming infection as long as 25 years after splenectomy. The splenectomized patient should be counseled about the risks of infection with certain organisms, such as the protozoan *Babesia* (Chap. 214) and *Capnocytophaga canimorsus* (formerly dysgonic fermenter 2 or DF-2), a bacterium carried in the mouths of animals (Chap. 127). Since encapsulated bacteria (*Streptococcus pneumoniae*, *Haemophilus influenzae*, and *Neisseria meningitidis*) are the organisms most commonly associated with postsplenectomy sepsis, splenectomized persons should be vaccinated (and revaccinated; Table 85-2) against the capsular polysaccharides of these organisms. Many clinicians recommend giving splenectomized patients a small supply of antibiotics effective against *S. pneumoniae*, *N. meningitidis*, and *H. influenzae* to avert rapid, overwhelming sepsis in the event that they cannot present for medical attention immediately after the onset of fever or other symptoms of bacterial infection.

The level of suspicion of infections with certain organisms should depend on the type of cancer diagnosed (Table 85-3). Diagnosis of multiple myeloma orCLLshould prompt the measurement of immunoglobulin levels and the consideration of either antibody

replacement or antibiotic prophylaxis. (In the case of CLL, antibiotic prophylaxis for likely pathogens has proven a cost-effective preventive measure.) Similarly, patients with acute lymphocytic leukemia (ALL), patients with non-Hodgkin's lymphoma, and all cancer patients treated with high-dose glucocorticoids (or glucocorticoid-containing chemotherapy regimens) should receive antibiotic prophylaxis for *Pneumocystis carinii* infection (Table 85-3).

In addition to exhibiting susceptibility to certain infectious organisms, patients with cancer are likely to manifest their infections in characteristic ways.

SYSTEM-SPECIFIC SYNDROMES

SKIN-SPECIFIC SYNDROMES (SEE PLATE IID-57)

Skin lesions are common in cancer patients, and their appearance may permit the diagnosis of systemic bacterial or fungal infection. While cellulitis caused by skin organisms such as *Streptococcus* or *Staphylococcus* is common, neutropenic patients and those with impaired blood or lymphatic drainage may develop infections with unusual organisms. Innocent-looking macules or papules may be the first sign of bacterial or fungal sepsis in immunocompromised patients. In the neutropenic host, a macule progresses rapidly to ecthyma gangrenosum (Fig. 19-CD1), a usually painless, round, necrotic lesion consisting of a central black or gray-black eschar with surrounding erythema. Ecthyma gangrenosum is located in nonpressure areas (as distinguished from necrotic lesions associated with lack of circulation) and is often associated with *Pseudomonas aeruginosa* bacteremia (Chap. 155) but may be caused by other bacteria.

Candidemia (Chap. 205) is also associated with a variety of skin conditions and commonly presents as a maculopapular rash. Punch biopsy of the skin may be the best method for diagnosis.

Cellulitis, an acute spreading inflammation of the skin, is most often caused by infection with group A *Streptococcus* or *Staphylococcus aureus*, virulent organisms normally found on the skin (Chap. 128). Although cellulitis tends to be circumscribed in normal hosts, it may spread rapidly in neutropenic patients [those with fewer than 500 functional polymorphonuclear leukocytes (PMNs) per microliter]. A tiny break in the skin may lead to spreading cellulitis, which is characterized by pain and erythema; in such patients, signs of infection (e.g., purulence) are often lacking. What might be a furuncle in a normal host may require amputation because of uncontrolled infection in a patient presenting with leukemia. A dramatic response to an infection that might be trivial in a normal host can mark the first sign of leukemia. Fortunately, granulocytopenic patients are likely to be infected with certain types of organisms (Table 85-4); thus the selection of an antibiotic regimen is somewhat easier than it might otherwise be. (See discussion below on the selection of antibiotics for use in neutropenic patients.) It is essential to recognize cellulitis early and to treat it aggressively. Patients who are neutropenic or have previously received antibiotics for other reasons may develop cellulitis with unusual organisms (e.g., *Escherichia coli*, *Pseudomonas*, or fungi). Early treatment, even of innocent-looking lesions, is essential to prevent necrosis and loss of tissue. Debridement to prevent spread may sometimes be necessary early in the course of

disease, but it can often be performed after chemotherapy, when thePMNcount increases.

Sweet's syndrome, or *febrile neutrophilic dermatosis*, was originally described in women with elevated white blood cell counts. The disease is characterized by the presence of leukocytes in the lower dermis, with edema of the papillary body. Ironically, this disease now is usually seen in neutropenic patients with cancer, most often in association with acute leukemia but also in association with a variety of other malignancies. Sweet's syndrome usually presents as red or bluish-red papules or nodules that may coalesce and form sharply bordered plaques. The edema may suggest vesicles, but on palpation the lesions are solid, and vesicles probably never arise in this disease. The lesions are most common on the face, neck, and arms. On the legs, they may be confused with erythema nodosum. The development of lesions is often accompanied by high fevers and an elevated erythrocyte sedimentation rate. Both the lesions and the temperature elevation respond dramatically to glucocorticoids. Treatment begins with high doses of glucocorticoids (60 mg of prednisone per day) followed by tapered doses over the next 2 to 3 weeks.

Data indicate that *erythema multiforme* with mucous membrane involvement is often associated with herpes simplex virus (HSV) infection and is distinct from Stevens-Johnson syndrome, which is associated with drugs and tends to have a more widespread distribution. Since cancer patients are both immunosuppressed (and therefore susceptible to herpes infections) and heavily treated with drugs (and therefore subject to Stevens-Johnson syndrome), both of these conditions are common in this population.

Cytokines, which are used as adjuvants or primary treatments for cancer, can themselves cause characteristic rashes, further complicating the differential diagnosis. This phenomenon is a particular problem in bone marrow transplant recipients (Chap. 136), who, in addition to having the usual chemotherapy-, antibiotic-, and cytokine-induced rashes, are plagued by graft-versus-host disease.

CATHETER-RELATED INFECTIONS

Because intravenous catheters are commonly used in cancer chemotherapy and are prone to infection (Chap. 135), they pose a major problem in the care of patients with cancer. Reviews have emphasized that some infected catheters can be treated with antibiotics while others must be removed. If the patient has a "tunneled" catheter (which consists of an entrance site, a subcutaneous tunnel, and an exit site), a red streak over the subcutaneous part of the line (the tunnel) is grounds for immediate removal of the catheter. Failure to remove catheters under these circumstances may result in extensive cellulitis and tissue necrosis.

More common than tunnel infections are exit-site infections, often with erythema around the area where the line penetrates the skin. Most authorities (Chap. 139) recommend treatment (usually with vancomycin) for an exit-site infection caused by a coagulase-negative *Staphylococcus*. Treatment of coagulase-positive staphylococcal infection is associated with a poorer outcome, and it is advisable to remove the catheter. Similarly, many clinicians remove catheters associated with infections due to *P.*

aeruginosa and *Candida* species, since such infections are difficult to treat and bloodstream infections with these organisms are likely to be deadly.

GASTROINTESTINAL TRACT-SPECIFIC SYNDROMES

Upper Gastrointestinal Tract Disease

Infections of the Mouth The oral cavity is rich in aerobic and anaerobic bacteria (Chap. 167) that normally live in a commensal relationship with the host. The antimetabolic effects of chemotherapy cause a breakdown of host defenses, leading to ulceration of the mouth and the potential for invasion by resident bacteria. Mouth ulcerations afflict most patients receiving chemotherapy and have been associated with viridans streptococcal bacteremia. A variety of topical rinses and elixirs have been proposed to treat these ulcerations. Although some may have a local anesthetic effect, the efficacy of any of these therapies in the prevention of disease is unproven. Similarly, the efficacy of mouthwashes in the prevention of esophagitis or invasive candidiasis is doubtful. Fluconazole, on the other hand, is clearly effective in the treatment of both local infections (thrush) and systemic infections (esophagitis) due to *C. albicans*.

Noma (or *cancrum oris*), commonly seen in malnourished children, is a penetrating disease of the soft and hard tissues of the mouth and adjacent sites, with resulting necrosis and gangrene. It has a counterpart in immunocompromised patients and is thought to be due to invasion of the tissues by *Bacteroides*, *Fusobacterium*, and other normal inhabitants of the mouth. It is associated with debility, poor oral hygiene, and immunosuppression.

Viruses, particularlyHSV, are a prominent cause of morbidity in immunocompromised patients, in whom they are associated with severe mucositis. The use of acyclovir, either prophylactically or therapeutically, is of value.

Esophageal Infections The differential diagnosis of esophagitis (usually presenting as substernal chest pain upon swallowing) includes herpes simplex and candidiasis, both of which are readily treatable.

Lower Gastrointestinal Tract Disease Hepatic candidiasis (Chap. 205) results from seeding of the liver (usually from a gastrointestinal source) in neutropenic patients. It is most common in patients being treated for acute leukemia and usually develops around the time the neutropenia resolves. The characteristic picture is that of persistent fever unresponsive to antibiotics; abdominal pain and tenderness or nausea; and elevated serum levels of alkaline phosphatase in a patient with hematologic malignancy who has recently recovered from neutropenia. The diagnosis of this disease (which may present in an indolent manner and persist for several months) is based on the finding of yeasts or pseudohyphae in granulomatous lesions. Hepatic ultrasound or computed tomography (CT) may reveal bull's-eye lesions. In some cases, magnetic resonance imaging (MRI) reveals small lesions not visible by other imaging modalities. The pathology (a granulomatous response) and the timing (with resolution of neutropenia and an elevation in granulocyte count) suggest that the host response to *Candida* is an important component of the manifestations of disease. In many cases, although organisms are visible, cultures of biopsied material may be negative. The designation

hepatosplenic candidiasis or *hepatic candidiasis* is a misnomer because the disease often involves the kidneys and other tissues; the term *chronic disseminated candidiasis* may be more appropriate. Because of the risk of bleeding with liver biopsy, diagnosis is often based on radiographic abnormalities. Amphotericin B is traditionally used for therapy (often for several months, until all manifestations of disease have disappeared), but fluconazole may be useful for outpatient therapy.

Typhlitis *Typhlitis*, sometimes referred to as necrotizing colitis, neutropenic colitis, necrotizing enteropathy, ileocecal syndrome, or cecitis, is a clinical syndrome of fever and right-lower-quadrant tenderness in an immunosuppressed host. This syndrome is almost always seen in neutropenic patients after chemotherapy with cytotoxic drugs. It may be more common among children than among adults and appears to be much more common among patients with acute myelocytic leukemia (AML) orALL than among those with other types of cancer. Physical examination reveals right-lower-quadrant tenderness, with or without rebound tenderness. Associated diarrhea (often bloody) is common, and the diagnosis can be confirmed by the finding of a thickened cecal wall onCT or ultrasonography. Plain films may reveal a right-lower-quadrant mass, but CT with contrast orMRI is a much more sensitive means of making the diagnosis. Although surgery is sometimes attempted to avoid perforation from ischemia, most cases resolve with medical therapy alone. The disease is sometimes associated with positive blood cultures (usually for aerobic gram-negative bacilli), and therapy is recommended for a broad spectrum of bacteria (particularly gram-negative bacilli, likely bowel flora). Recurrence is rare, and most patients recover uneventfully.

***Clostridium difficile*-Induced Diarrhea** Cancer patients seem to be predisposed to the development of *C. difficile* diarrhea (Chap. 145) as a consequence of chemotherapy alone. Thus, they may have positive toxin tests before receiving antibiotics. Obviously, such patients are also subject to *C. difficile*-induced diarrhea as a result of antibiotic pressure. It is worth noting that toxins other than *C. difficile* may be associated with diarrhea; therefore, the detection of nonspecific toxins in the stool -- without a specific neutralization test -- does not prove that *C. difficile* infection is present.

CENTRAL NERVOUS SYSTEM-SPECIFIC SYNDROMES

Meningitis While meningitis in immunocompetent adults is likely to be caused by *S. pneumoniae*, the same is not true in immunocompromised patients. As noted previously, splenectomized patients are susceptible to rapid overwhelming infection with encapsulated bacteria (including *S. pneumoniae*, *H. influenzae*, and *N. meningitidis*). Similarly, patients who are antibody-deficient (such as patients withCLL, those who have received intensive chemotherapy, or those who have undergone bone marrow transplantation) are likely to have infections with these bacteria. Other cancer patients, however, because of their defective cellular immunity, are likely to be infected with other pathogens (Table 85-3). The presentation of meningitis in patients with lymphoma, patients receiving chemotherapy (particularly with glucocorticoids) for solid tumors, and patients who have received bone marrow transplants suggests a diagnosis of cryptococcal or listerial infection.

Encephalitis The spectrum of disease resulting from viral encephalitis is expanded in immunocompromised patients. Infection with varicella-zoster virus (VZV) has been

associated with encephalitis that may be caused by VZV-related vasculitis. The slow viruses (e.g., Creutzfeldt-Jakob agent) may also be associated with dementia and encephalitic presentations, and a diagnosis of progressive multifocal leukoencephalopathy should be considered when a patient who has received chemotherapy presents with dementia. Other abnormalities of the central nervous system (CNS) that may be confused with infection include normal-pressure hydrocephalus and vasculitis resulting from CNS irradiation. It may be possible to differentiate these conditions byMRI.

Brain Abscess Brain abscesses in immunocompromised patients are likely to be due to *Cryptococcus* (particularly in patients with lymphoma or those receiving glucocorticoids), *Nocardia*, or *Aspergillus*. *Aspergillus* may enter via the lungs or -- like *Mucor* -- may invade the hard and soft palates to cause pneumonia (see below) with or without brain abscesses.

PULMONARY INFECTIONS

Pneumonia (Chap. 255) in immunocompromised patients may be difficult to diagnose because conventional methods of diagnosis depend on the presence of neutrophils. Bacterial pneumonia in neutropenic patients may present without purulent sputum -- or, in fact, without any sputum at all -- and may not produce physical findings suggestive of chest consolidation (rales or egophony).

In granulocytopenic patients with persistent or recurrent fever, the chest x-ray pattern may help to localize an infection and thus to determine which investigative tests and procedures should be undertaken and which therapeutic options should be considered (Table 85-5). The difficulties encountered in the management of pulmonary infiltrates relate in part to the difficulties of performing diagnostic procedures on the patients involved. When platelet counts can be increased to adequate levels by transfusion, microscopic and microbiologic evaluation of the fluid obtained by endoscopic bronchial lavage is often diagnostic. Lavage fluid should be cultured for *Mycoplasma*, *Chlamydia*, *Legionella*, *Nocardia*, fungi, and more common bacterial pathogens. In addition, the possibility of *P. carinii* pneumonia should be considered, especially in patients withALL or lymphoma who have not received prophylactic trimethoprim-sulfamethoxazole. The characteristics of the infiltrate may be helpful in decisions about further diagnostic and therapeutic maneuvers. Nodular infiltrates suggest fungal pneumonia (e.g., that caused by *Aspergillus* or *Mucor*). Such lesions may best be approached by visualized biopsy procedures.

Aspergillus species (Chap. 206) can colonize the skin and respiratory tract or cause fatal systemic illness. Although *Aspergillus* may cause aspergillomas in a previously existing cavity or may produce allergic bronchopulmonary aspergillosis, the major problem posed by this genus in neutropenic patients is invasive disease due to *A. fumigatus* or *A. flavus*. The organisms enter the host through colonization of the respiratory tract, with subsequent invasion of the blood vessels. The disease is likely to present as a thrombotic or embolic event because of the ability of the organisms to invade blood vessels. The risk of infection with *Aspergillus* correlates directly with the duration of neutropenia. In prolonged neutropenia, positive surveillance cultures for colonization of the nasopharynx with *Aspergillus* may predict the development of

disease.

Patients with *Aspergillus* infection often present with pleuritic chest pain and fever, which are sometimes accompanied by cough. Hemoptysis may be an ominous sign. Chest x-rays may reveal new focal infiltrates or nodules. ChestCT may reveal a characteristic halo consisting of a mass-like infiltrate surrounded by an area of low attenuation. The presence of a "crescent sign" on a chest x-ray or a chest CT scan, in which the mass progresses to central cavitation, is characteristic of invasive *Aspergillus* infection but may develop only with the resolution of the lesions.

In addition to causing pulmonary presentations, *Aspergillus* may invade through the nose or palate, with deep sinus penetration. The appearance of a discolored area in the nasal passages or on the hard palate should prompt a search for invasive *Aspergillus*. This situation is likely to require surgical debridement. Treatment (Chap. 206) with high doses of amphotericin B has been successful in curing granulocytopenic patients of invasive *Aspergillus* infection after the return of granulocytes. Catheter infections with *Aspergillus* usually require both removal of the catheter and antifungal therapy.

Diffuse interstitial infiltrates suggest viral, parasitic, or *P. carinii* pneumonia. If the patient has a diffuse interstitial pattern on chest x-ray, it may be reasonable to institute empirical treatment with trimethoprim-sulfamethoxazole (for *Pneumocystis*) and an erythromycin derivative or a quinolone (for *Chlamydia*, *Mycoplasma*, and *Legionella*) while considering invasive diagnostic procedures. Noninvasive procedures, such as staining of sputum smears for *Pneumocystis* and serum cryptococcal antigen tests, may be helpful on occasion. In transplant recipients who are seropositive for cytomegalovirus (CMV), culture of a nonpulmonary site for CMV may be worthwhile. Infections with viruses that cause only upper respiratory symptoms in immunocompetent hosts, such as respiratory syncytial, influenza, and parainfluenza viruses, may be associated with fatal pneumonitis in immunocompromised hosts. An attempt at early diagnosis by nasopharyngeal aspiration should be considered so that appropriate treatment can be instituted.

While bleomycin is the most common cause of chemotherapy-induced lung disease, other causes include alkylating agents (such as cyclophosphamide, chlorambucil, and melphalan), nitrosoureas [carmustine (BCNU), lomustine (CCNU), and methyl-CCNU], busulfan, procarbazine, methotrexate, and hydroxyurea. Both infectious and noninfectious (drug- and/or radiation-induced) pneumonitis can cause fever and abnormalities on chest x-ray; thus, the differential diagnosis of an infiltrate in a patient receiving chemotherapy encompasses a broad range of conditions (Table 85-5). Since the treatment of radiation pneumonitis (which may respond dramatically to glucocorticoids) or drug-induced pneumonitis is different from that of infectious pneumonia, a biopsy may be important in the diagnosis. Unfortunately, no definitive diagnosis can be made in approximately 30% of cases, even after bronchoscopy.

Open-lung biopsy is the "gold standard" of diagnostic techniques. Biopsy via a visualized thoracostomy can replace an open procedure in many cases. When a biopsy cannot be performed, empirical treatment can be undertaken with erythromycin (or an erythromycin derivative such as azithromycin) and trimethoprim-sulfamethoxazole (in the case of diffuse infiltrates) or with amphotericin B (in the case of nodular infiltrates).

The risks should be weighed carefully in these cases. If inappropriate drugs are administered, empirical treatment may prove toxic or ineffective; either of these outcomes may be riskier than biopsy.

CARDIOVASCULAR INFECTIONS

Patients with Hodgkin's disease are prone to persistent infections by *Salmonella*, sometimes (and particularly often in elderly patients) affecting a vascular site. The use of intravenous catheters deliberately lodged in the right atrium is associated with a high incidence of bacterial endocarditis (presumably related to valve damage followed by bacteremia). Nonbacterial thrombotic endocarditis has been described in association with a variety of malignancies (most often solid tumors) and may follow bone marrow transplantation as well. The presentation of an embolic event with a new cardiac murmur suggests this diagnosis. Blood cultures are negative in this disease of unknown pathogenesis.

ENDOCRINE SYNDROMES

In addition to infections of the skin, gastrointestinal tract, and pulmonary system, infections of the endocrine system have been described in immunocompromised patients. *Candida* infection of the thyroid during neutropenia can be defined by indium-labeled white-cell scans or gallium scans after neutrophil counts increase.CMVinfection can cause adrenalitis with or without resulting adrenal insufficiency. The presentation of a sudden endocrine anomaly in an immunocompromised patient may be a sign of infection in the involved end organ.

MUSCULOSKELETAL INFECTIONS

Infection that is a result of vascular compromise (resulting in gangrene) can occur when a tumor compromises the blood supply to muscles, bones, or joints. The process of diagnosis and treatment of such infection is similar to that in normal hosts, with the following caveats: (1) In terms of diagnosis, a lack of physical findings resulting from a lack of granulocytes in the granulocytopenic patient should make the clinician more aggressive in obtaining tissue rather than relying on physical signs. (2) In terms of therapy, aggressive debridement of infected tissues may be required, but it is usually difficult to operate on patients who have recently received chemotherapy, both because of a lack of platelets (which results in bleeding complications) and because of a lack of white blood cells (which may lead to secondary infection). A blood culture positive for *Clostridium perfringens* (an organism commonly associated with gas gangrene) can have a number of meanings (Chap. 145). Bloodstream infections with intestinal organisms like *Streptococcus bovis* and *C. perfringens* may arise spontaneously from lower gastrointestinal lesions (tumor or polyps); alternatively, these lesions may be harbingers of invasive disease. The clinical setting must be considered in order to define the appropriate treatment for each case.

RENAL AND URETERAL INFECTIONS

Infections of the urinary tract are common among patients whose ureteral excretion is compromised (Table 85-1). *Candida*, which has a predilection for the kidney, can invade

either from the bloodstream or in a retrograde manner (via the ureters or bladder) in immunocompromised patients. The presence of "fungus balls" or persistent candiduria suggests invasive disease. Persistent funguria (with *Aspergillus* as well as *Candida*) should prompt a search for a nidus of infection in the kidney.

Certain viruses are typically seen only in immunosuppression. BK virus (polyomavirus hominis 1) has been documented in the urine of bone marrow transplant recipients and, like adenovirus, may be associated with hemorrhagic cystitis. BK-induced cystitis usually remits with decreasing immunosuppression. Anecdotal reports have described the treatment of adenovirus with ribavirin in cases of severe hemorrhagic cystitis in immunocompromised patients.

ABNORMALITIES THAT PREDISPOSE TO INFECTION

THE LYMPHOID SYSTEM

It is beyond the scope of this chapter to detail how all the immunologic abnormalities that result from cancer or from chemotherapy for cancer lead to infections. Disorders of the immune system are discussed in other sections of this book. As has been noted, patients with antibody deficiency are predisposed to overwhelming infection with encapsulated bacteria (including *S. pneumoniae*, *H. influenzae*, and *N. meningitidis*). Infections that result from the lack of a functional cellular immune system are described inChap. 309. It is worth mentioning, however, that patients undergoing intensive chemotherapy for any form of cancer will have not only defects due to granulocytopenia but also lymphocyte dysfunction, which may be profound. Thus, these patients -- especially those receiving glucocorticoid-containing regimens -- should be given prophylaxis for *P. carinii* pneumonia.

THE HEMATOPOIETIC SYSTEM

Initial studies in the 1960s revealed a dramatic increase in the incidence of infections (fatal and nonfatal) among cancer patients with a granulocyte count of<500/uL. Recent studies have cited a figure of 48.3 infections per 100 neutropenic patients (<1000 granulocytes per microliter) with hematologic malignancies and solid tumors, or 46.3 infections per 1000 days at risk.

Neutropenic patients are unusually susceptible to infection with a wide variety of bacteria; thus, antibiotic therapy should be initiated promptly to cover likely pathogens if infection is suspected. Indeed, early initiation of antibacterial agents is mandatory to prevent deaths. These patients are susceptible to gram-positive and gram-negative organisms found commonly on the skin and in the bowel (Table 85-4). Because treatment with narrow-spectrum agents leads to infection with organisms not covered by the antibiotics used, the initial regimen should target pathogens likely to be initial causes of bacterial infection in neutropenic hosts (Fig. 85-1).

TREATMENT

Antibacterial Therapy Hundreds of antibacterial regimens have been tested for use in neutropenic patients with cancer. Many of the relevant studies involved small

populations in which the outcomes were generally good, and most lacked the statistical power to detect differences among the regimens studied. Each febrile neutropenic patient should be approached as a unique problem, with particular attention given to previous infections and recent exposures to antibiotics. Several general guidelines are useful in the initial treatment of neutropenic patients with fever (Fig. 85-1):

1. It is necessary to use antibiotics active against both gram-negative and gram-positive bacteria (Table 85-4) in the initial regimen.

2. An aminoglycoside or an antibiotic without good activity against gram-positive organisms (e.g., ciprofloxacin) alone is not adequate in this setting.

3. The agents used should reflect both the epidemiology and the antibiotic resistance pattern of the hospital. For example, in hospitals where there is gentamicin resistance, amikacin-containing regimens should be considered; in hospitals with frequent *P. aeruginosa* infections, a regimen with the highest level of activity against this pathogen (such as tobramycin plus a semisynthetic penicillin) would be reasonable for initial therapy.

4. A single third-generation cephalosporin constitutes an appropriate initial regimen in many hospitals (if the pattern of resistance justifies its use).

5. Most standard regimens are designed for patients who have not previously received prophylactic antibiotics. The development of fever in a patient receiving antibiotics affects the choice of subsequent therapy (which should target resistant organisms and organisms known to cause infections in patients being treated with the antibiotics already administered).

6. Randomized trials have indicated that it is safe to use oral antibiotic regimens to treat "low-risk" patients with fever and neutropenia. Outpatients who are expected to remain neutropenic for <10 days and who have no concurrent medical problems (such as hypotension, pulmonary compromise, or abdominal pain) can be classified as low risk and treated with a broad-spectrum oral regimen. On the basis of large studies, it can be concluded that this therapy is safe and effective, at least when delivered in the inpatient setting. Outpatient treatment has been assessed in small studies, but data from large randomized trials demonstrating the safety of outpatient treatment of fever and neutropenia are not yet available.

The initial antibacterial regimen should be refined on the basis of culture results (Fig. 85-1). Blood cultures are the most relevant on which to base therapy; surface cultures of skin and mucous membranes may be misleading. In the case of gram-positive bacteremia or another gram-positive infection, it is important that the antibiotic be optimal for the organism isolated. If the infection is caused by certain gram-negative pathogens (such as *P. aeruginosa*), a synergistic combination of antibiotics (usually a semisynthetic penicillin, such as piperacillin, plus an aminoglycoside) may be appropriate. Although it is not desirable to leave the patient unprotected, the addition of more and more antibacterial agents to the regimen is not appropriate unless there is a clinical or microbiologic reason to do so. *Planned progressive therapy* (the serial, empirical addition of one drug after another without culture data) is not efficacious in

most settings and may have unfortunate consequences. Simply adding another antibiotic for fear that a gram-negative infection is present is a dubious practice. The synergy exhibited by b-lactams and aminoglycosides against certain gram-negative organisms (especially *P. aeruginosa*) provides the rationale for using two antibiotics in this setting. Mere addition of a quinolone or another antibiotic not likely to exhibit synergy for "double coverage" has not been shown to be of benefit and may cause additional toxicities and side effects. Cephalosporins can cause bone marrow suppression, and vancomycin is associated with neutropenia in some healthy people (Chap. 137). Furthermore, the addition of multiple cephalosporins may induce b-lactamase production by some organisms; cephalosporins and double b-lactam combinations should probably be avoided altogether in *Enterobacter* infections.

Antifungal Therapy Fungal infections in cancer patients are most often associated with neutropenia. Neutropenic patients are predisposed to the development of invasive fungal infections, most commonly those due to *Candida* and *Aspergillus* species and occasionally those caused by *Fusarium*, *Trichosporon*, and *Bipolaris*. Cryptococcal infection, which is common among patients taking immunosuppressive agents, is uncommon among neutropenic patients receiving chemotherapy forAML. Invasive candidal disease is usually caused by *C. albicans* or *C. tropicalis* but can be caused by *C. krusei*, *C. parapsilosis*, and *C. glabrata*.

Most clinicians add amphotericin B to antibacterial regimens if a neutropenic patient remains febrile despite 4 to 7 days of treatment with antibacterial agents. The rationale for the empirical addition of amphotericin B is that it is difficult to culture fungi before they cause disseminated disease and that mortality from disseminated fungal infections in granulocytopenic patients is high. The imidazoles (especially fluconazole) may have prophylactic efficacy in this regard, but the spectrum of activity of the currently available azoles is narrower than that of amphotericin B. Amphotericin B is the mainstay of therapy for disseminated *Candida* or *Aspergillus* infection in the neutropenic patient. The combined use of an imidazole and amphotericin B is controversial because of the theoretical antagonistic effects of these agents. The insolubility of amphotericin B has resulted in the marketing of several amphotericin B-lipid formulations. Lipid preparations have been shown to be less toxic than the amphotericin B deoxycholate complex. However, because of the high cost of the lipid preparations, at many centers their use is reserved for patients who fail to respond to standard amphotericin B. Since the side effects of the formulations differ, unnecessary switching from one to another is not recommended.

Other Therapeutic Modalities Another way to address the problems of the febrile neutropenic patient is to replenish the neutrophil population. Although granulocyte transfusions are efficacious in the treatment of refractory gram-negative bacteremia, they do not have a documented role in prophylaxis. Because of the expense, the risk of leukoagglutinin reactions (although this risk has probably been decreased by improved cell-separation procedures), and the risk of transmission ofCMV from unscreened donors, granulocyte transfusion is reserved for patients unresponsive to antibiotics. This modality is efficacious for documented gram-negative bacteremia refractory to antibiotics, particularly in situations where granulocyte numbers will be depressed for only a short period.

A variety of cytokines, including granulocyte colony-stimulating factor and granulocyte-macrophage colony-stimulating factor, enhance granulocyte recovery after chemotherapy and consequently shorten the period of maximal vulnerability to fatal infections. The role of these cytokines in routine practice is still a matter of some debate. Most authorities recommend their use only when neutropenia is both severe and prolonged. The cytokines themselves may have adverse effects, including fever, hypoxemia, and pleural effusions or serositis in other areas. Since there is little evidence that their routine administration lessens the risk of death and since they are still expensive, the cytokines have not become the standard of care in all centers. The role of other cytokines (such as macrophage colony-stimulating factor for monocytes or interferon-g) in preventing or treating infections in granulocytopenic patients is under investigation.

Once neutropenia has resolved, patients are not at high risk of infection. However, depending on what drugs they receive, patients who continue on chemotherapeutic protocols remain at high risk for certain diseases. Any patient receiving more than a maintenance dose of glucocorticoids (including many treatment regimens for diffuse lymphoma) should also receive prophylactic trimethoprim-sulfamethoxazole because of the risk of *P. carinii* infection; those withALLshould receive such prophylaxis for the duration of chemotherapy.

PREVENTION OF INFECTION IN CANCER PATIENTS

EFFECT OF THE ENVIRONMENT

Outbreaks of fatal *Aspergillus* infection have been associated with construction projects and materials in several hospitals. The association between spore counts and risk of infection suggests the need for a high-efficiency air-handling system in hospitals that care for large numbers of neutropenic patients. The use of laminar-flow rooms and prophylactic antibiotics has decreased the number of infectious episodes in severely neutropenic patients. However, because of the expense of such a program and the failure to show that it dramatically affects mortality, most centers do not routinely use laminar flow to care for neutropenic patients. Some centers use "reverse isolation," in which health care providers and visitors to a patient who is neutropenic wear gowns and gloves. Since most of the infections these patients develop are due to organisms that colonize the patients' own skin and bowel, the validity of such schemes is dubious, and limited clinical data do not support their use. Hand washing by all staff caring for neutropenic patients should be required to prevent the spread of resistant organisms.

The presence of large numbers of bacteria (particularly *P. aeruginosa*) in certain foods, especially fresh vegetables, has led some authorities to recommend a special "low-bacteria" diet. A diet consisting of cooked and canned food is satisfactory to most neutropenic patients and does not involve elaborate disinfection or sterilization protocols. However, there are no studies to support even this type of dietary restriction. Counseling of patients to avoid leftovers, deli foods, and unpasteurized dairy products is recommended.

PHYSICAL MEASURES

Although few studies address this issue, patients with cancer are predisposed to infections resulting from anatomic compromise (e.g., lymphedema resulting from node dissections after radical mastectomy). Surgeons who specialize in cancer surgery can provide specific guidelines for the care of such patients, and patients benefit from common-sense advice about how to prevent infections in vulnerable areas.

ANTIBIOTIC PROPHYLAXIS

There is no consensus on the use of prophylactic antibiotics in neutropenic patients. The incidence of infection is lower among patients who receive broad-spectrum antibiotic prophylaxis than among those who do not. Because of the prolongation of neutropenia associated with the use of trimethoprim-sulfamethoxazole, some clinicians use broad-spectrum agents such as quinolones (e.g., ciprofloxacin). Either regimen can be given orally, and both have the advantage of inactivity against anaerobic organisms; thus, neither is likely to disrupt the bowel flora and permit colonization with new aerobes or *Candida*. However, both regimens have adverse effects and can lead to the selection of resistant organisms in a hospital. For these reasons, many clinicians reserve their use for patients with the longest periods of neutropenia (e.g., bone marrow transplant recipients). The same issues apply to the use of antifungal agents. While agents such as fluconazole may prevent infections with susceptible organisms (e.g., *C. albicans*), they can cause a concomitant increase in infections due to resistant fungi (e.g., *C. krusei*). Thus, the decision to use antifungal prophylaxis may vary with the fungi endemic in a given hospital. Prophylaxis for *P. carinii* is mandatory for patients withALL and for all cancer patients receiving glucocorticoid-containing chemotherapy regimens.

VACCINATION OF CANCER PATIENTS

In general, patients undergoing chemotherapy respond less well to vaccines than normal hosts. Their greater need for vaccines thus leads to a dilemma in their management. Purified proteins and inactivated vaccines are almost never contraindicated and should be given to patients even during chemotherapy. For example, all adults should receive diphtheria-tetanus toxoid boosters at the indicated times as well as seasonal influenza vaccine. However, if possible, vaccination should not be undertaken concurrent with cytotoxic chemotherapy. If patients are expected to be receiving chemotherapy for several months and vaccination is indicated (for example, influenza vaccination in the fall), the vaccine should be given midcycle -- as far as possible from the antimetabolic agents that will prevent an immune response. The meningococcal and pneumococcal polysaccharide vaccines should be given to patients before splenectomy, if possible. The Advisory Committee on Immunization Practices recommends reimmunization every 5 years for the pneumococcal vaccine; although no official stand has been taken, this recommendation seems reasonable for the meningococcal vaccine as well. The *H. influenzae* type b conjugate vaccine should be administered to all splenectomized patients; there is no current recommendation for reimmunization, but immunity appears to be much longer-lasting than that induced by polysaccharide vaccines.

In general, live virus (or live bacterial) vaccines should not be given to patients during intensive chemotherapy because of the risk of disseminated infection. Recommendations on vaccination are summarized in Table 85-2.

(Bibliography omitted in Palm version)

Back to Table of Contents

86. MELANOMA AND OTHER SKIN CANCERS - *Arthur J. Sober*, *Howard K. Koh*, *Gregory P. Wittenberg*, *Carl V. Washington, Jr.*

Pigmented skin lesions are among the most common findings on physical examination. The challenge is to distinguish cutaneous melanomas, which may be lethal, from the remainder, which with rare exceptions are benign. Cutaneous neoplasms are depicted in Section IIB (Plates IIB-20,IIB-21,IIB-22,IIB-23,IIB-24,IIB-25,IIB-26, andIIB-27) of the Color Atlas; benign and malignant pigmented lesions are in Section IIC (Plates IIC-28,IIC-29,IIC-30,IIC-31,IIC-31, andIIC-33).

MELANOMA

Melanomas originate from melanocytes, pigment cells normally present in the epidermis and sometimes in the dermis. This tumor affects approximately 44,200 individuals per year in the United States, resulting in 7300 deaths. The tumor can affect adults of all ages, even teenagers; it has distinct clinical features that make it detectable at a time when cure by surgical excision is possible; and it is located on the skin surface, where it is visible. The incidence has increased dramatically; if the incidence continues to increase at the present rate, within a decade, lifetime risk of melanoma will exceed 1%.

The reason for the increase in melanoma incidence is thought to be increased recreational sun exposure, especially early in life. Individuals of similar ethnic background who immigrate after childhood to areas of high sun exposure (e.g., Israel and Australia) have lower melanoma rates than individuals of similar age who were either born in those countries or immigrated before age 10. The individuals most susceptible to development of melanoma are those with fair complexions, red or blond hair, blue eyes, and freckles and who tan poorly and sunburn easily. Most studies link increased melanoma risk to history of sunburn. Other factors associated with increased risk include a family history of melanoma (1 in 10 melanoma patients have an affected family member); the presence of a clinically atypical mole (dysplastic nevus), a giant congenital melanocytic nevus, or a small to medium-sized congenital melanocytic nevus (see below); the presence of a higher than average number of ordinary melanocytic nevi; and immunosuppression (Table 86-1). Individuals with 50 or more moles³2 mm in size have a 64-fold increased risk. About 30% of melanomas arise in a nevus. Melanoma is relatively rare in heavily pigmented peoples. Dark-skinned populations (such as those of India and Puerto Rico), blacks, and East Asians have rates 10 to 20 times lower than lighter-skinned whites. In keeping with the role of sun exposure, the incidence is inversely correlated with the latitude of residence; at any latitude, however, darker-skinned persons have a lower incidence.

CLINICAL CHARACTERISTICS

There are four types of cutaneous melanoma (Table 86-2). In three of these -- superficial spreading melanoma, lentigo maligna melanoma, and acral lentiginous melanoma -- the lesion has a period of superficial (so-called radial) growth during which it increases in size but does not penetrate deeply. During this period, the melanoma is most capable of being cured by surgical excision. The fourth type -- nodular melanoma -- does not have a recognizable radial growth phase and usually presents as a deeply invasive lesion, capable of early metastasis. When tumors begin to penetrate deeply

into the skin, they are in the so-called vertical growth phase. Melanomas with a radial growth phase are characterized by irregular and sometimes notched borders, variation in pigment pattern, and variation in color. An increase in size or change in color is noted by the patient in 70% of early lesions. Bleeding, ulceration, and pain are late signs and are of little help in early recognition. Nodular melanomas are dark brown-black to blue-black nodules. Melanomas are occasionally amelanotic, in which case the diagnosis is established histologically after biopsy of a new or changing skin growth. Lentigo maligna melanoma is usually confined to chronically sun-damaged sites (face, neck, back of hands) in older individuals. Acral lentiginous melanoma occurs on the palms, soles, nail beds, and mucous membranes. While this type occurs in whites, it is most frequent (along with nodular melanoma) in blacks and East Asians. Superficial spreading melanoma is most frequent in whites. Melanomas arising in dysplastic nevi (see below) are usually of this type. The back is the most common site for melanoma in men. In women, the back and the lower leg (knee to ankle) are frequent sites.

PROGNOSTIC FACTORS

The most important prognostic factor is the stage at the time of presentation [see the discussion of revised American Joint Cancer Commission (AJCC) stages, below]. Five-year survival for clinical stages I and II (primary tumor; no clinical evidence of disease elsewhere) is about 85%. For clinical stage III (clinically palpable regional nodes that contain tumor), the 5-year survival is about 50% when only one node is involved and about 15 to 20% when four or more nodes are involved. The 5-year survival for clinical stage IV (disseminated disease) is<5%. Fortunately, most melanomas are diagnosed in clinical stages I and II. Within these stages, the prognosis depends on the thickness of the primary tumor (Table 86-3). This system is based on the rationale that the likelihood of metastasis should correlate with tumor volume, with thickness being the best single index of tumor volume. Melanomas<0.76 mm thick are usually cured by surgical removal (with 5-year survival rates of 96 to 99%). Approximately 40% of primary melanomas now fall in a low-risk category (thickness<1 mm). In low-risk patients who develop metastases, the primary tumors often exhibit either microscopic features of anaplasia or a vertical growth phase. More than 50% of individuals with melanomas ³4 mm thick will develop metastatic disease and die of their melanoma (Table 86-3). These thick tumors are almost always raised above the plane of the skin. Certain anatomic sites affect the prognosis. The favorable sites appear to be the forearm and leg (excluding feet), while unfavorable sites include scalp, hands, feet, and mucous membranes. In general, women with stage I or II disease have a better survival than men, perhaps in part because of earlier diagnosis; women frequently have melanomas on the lower leg, where self-recognition is more likely and prognosis is better. Older individuals have poorer prognoses. This finding has been explained in part by a tendency toward later diagnosis (and thus thicker tumors) in men and by a higher proportion in men of acral melanomas (palmar-plantar), which have a poorer prognosis. Melanoma may recur after many years. About 10 to 15% of first-time recurrences develop more than 5 years after treatment of the original lesion. The time to recurrence varies inversely with tumor thickness. Other prognostic factors for stages I and II melanoma include the presence of an ulcer in the primary tumor, high mitotic rate, and the presence of microscopic tumor satellites (foci of tumor ³0.05 mm in diameter in the reticular dermis or subcutaneous fat, distinct from the main body of the tumor). The presence of microscopic satellites is also predictive of microscopic metastases to the

regional lymph nodes. An alternative prognostic scheme for clinical stages I and II melanoma, proposed by Clark, is based on the anatomic level of invasion in the skin. Level I is intraepidermal (in situ); level II penetrates the papillary dermis; level III spans the papillary dermis; level IV penetrates the reticular dermis; and level V penetrates into the subcutaneous fat. The 5-year survival for these stages averages 100, 95, 82, 71, and 49%, respectively.

NATURAL HISTORY

Melanomas may spread by the lymphatic channels or the bloodstream. The earliest metastases are often to regional lymph nodes. Surgical lymphadenectomy usually controls regional disease. Liver, lung, bone, and brain are common sites of hematogenous spread, but unusual sites, such as the anterior chamber of the eye, may also be involved. Once metastatic disease is established, the likelihood of cure is low.

MANAGEMENT

The entire cutaneous surface, including the scalp and mucous membranes, should be examined in each patient. Bright room illumination is important, and a 7´ to 10´ hand lens is helpful for evaluating variation in pigment pattern. A history of relevant risk factors should be elicited. Any suspicious lesions should be biopsied, evaluated by a specialist, or recorded by chart and/or photography for follow-up. Examination of the lymph nodes and palpation of the abdominal viscera are part of the staging examination for suspected melanoma. The patient should be advised to have other family members screened if either melanoma or clinically atypical moles (dysplastic nevi) are present. Melanoma prevention is based on protection from the sun. Routine use of a sunblock of SPF³15, use of protective clothing, and avoiding intense midday ultraviolet exposure should be recommended. The patient should be educated in the clinical features of melanoma and advised to report any growth or other change in a pigmented lesion. Patient education brochures are available from the American Cancer Society, the American Academy of Dermatology, the National Cancer Institute, and the Skin Cancer Foundation. Self-examination at 6- to 8-week intervals may enhance the likelihood of detecting change between follow-up visits. Routine follow-up visits for melanoma patients and patients with clinically atypical moles (dysplastic nevi) may facilitate early detection of new tumors.

Precursor Lesions Clinically atypical moles, also termed *dysplastic nevi*, occur in certain families affected by melanoma. In some families, melanomas occur nearly exclusively in the individuals with dysplastic nevi. These nevi appear to be transmitted as an autosomal dominant trait that involves chromosome 9p16. In other families, the nevi may not be present in all individuals with an increased risk of melanoma. The melanomas may arise in clinically atypical moles or in normal skin. Individuals with clinically atypical moles and two family members with melanoma have been reported to have a >50% lifetime risk for developing melanoma.Table 86-4 lists the features that are characteristic of clinically atypical moles and that differentiate them from benign acquired nevi. The number of clinically atypical moles may vary from one to several hundred. Clinically atypical moles usually differ from each other in appearance. The borders are often hazy and indistinct, and the pigment pattern is more highly varied than that in benign acquired nevi. Of the 90% of melanoma patients whose disease is

regarded as sporadic (i.e., who lack a family history of melanoma), about 40% have clinically atypical moles, as compared with an estimated 5 to 10% of the population at large. The observation that at least 20% of sporadic melanomas arise in association with a clinically atypical mole makes this nevus the most important precursor for melanoma.

Less frequent precursors include the giant congenital melanocytic nevus and the small congenital melanocytic nevus (although the latter relationship is disputed by some). Congenital melanocytic nevi are present at birth or appear in the neonatal period (tardive form). The giant melanocytic nevus, also called the bathing trunk, cape, or garment nevus, is a rare malformation that affects perhaps 1 in 30,000 to 1 in 100,000 individuals. These nevi are usually>20 cm in diameter and may cover more than half the body surface. Giant nevi often occur in association with multiple small congenital nevi. The borders are sharp, and hair may be present. The lesions are usually dark brown and may have darker and lighter areas. Pigment is haphazardly displayed. The surface is smooth to rugose or cerebriform and may vary from one portion of the lesion to another. A lifetime risk of melanoma development of 6% has been estimated. The risk is greatest before age 5 and next greatest between ages 5 and 10. Early detection of melanoma is difficult in these lesions because of the deep dermal or subcutaneous origin of melanoma in these lesions and because of the large and varied surface of the nevus. Prophylactic excision early in life can be accomplished by staged removal with coverage by split-thickness skin grafts. No uniform management guidelines for giant congenital nevi have been developed.

The small- to medium-sized congenital melanocytic nevus, which affects approximately 1% of persons, usually presents as a raised dark- to medium-brown lesion with a smooth or papillomatous surface. The border is sharp, and lesions may be oriented along lines of skin cleavage. Follicular hyper- and hypopigmentation may coexist in a salt-and-pepper configuration. The lesion may have an excess of coarse hairs. Melanoma may develop in these lesions but the risk is not quantitated. Considerations of body surface area suggest that the incidence of melanomas arising in small congenital melanocytic nevi is probably higher than would be expected by chance. The remnants of a nevus with histopathologic features of a congenital nevus have been observed in 2 to 6% of melanomas. The management of small- to medium-sized congenital melanocytic nevi remains controversial; prophylactic removal under local anesthesia in the early teen years is appropriate as melanomas arise later in these lesions.

Differential Diagnosis The aim of differential diagnosis is to distinguish benign pigmented lesions from melanoma and its precursors. If melanoma is a consideration, then biopsy is appropriate. Some benign look-alikes may be removed in the process of trying to detect authentic melanoma. Table 86-5 summarizes the distinguishing features of benign lesions that may be confused with melanoma. Early detection of melanoma may be facilitated by applying the "ABCD rules": A -- asymmetry, benign lesions are usually symmetric; B -- border irregularity, most nevi have clear-cut borders; C -- color variegation, benign lesions usually have uniform light or dark pigment; D -- diameter >6 mm (the size of a pencil eraser).

Biopsy Any pigmented cutaneous lesion that has changed in size or shape or has other

features suggestive of malignant melanoma should be biopsied. The recommended technique is an excisional biopsy, as that facilitates pathologic assessment of the lesion, permits accurate measurement of thickness if the lesion is melanoma, and constitutes treatment if the lesion is benign. Shave biopsy or curettage of a suspected melanoma is contraindicated. For large lesions or lesions on anatomic sites where excisional biopsy may not be feasible (such as the face, hands, or feet), an incisional biopsy through the most nodular or darkest area of the lesion is acceptable; this should include the vertical growth phase of the primary tumor, if present. Data from prospective studies do not indicate that an incisional biopsy facilitates the spread of melanoma.

Staging Once the diagnosis of malignant melanoma has been confirmed, the tumor must be staged to determine prognosis and treatment. The history should probe for evidence of metastatic disease, such as malaise, weight loss, headaches, visual difficulty, or bone pain. The physical examination should be directed especially to the skin, regional draining lymph nodes, central nervous system, liver, and spleen. In the absence of signs or symptoms of metastasis, few laboratory or radiologic tests are indicated for staging purposes. Aside from a chest radiograph, and possibly liver function tests, no other tests or scans are routinely indicated unless the history or physical examination suggests metastasis to a specific organ. Specifically, liver-spleen scans and computed tomography have a low yield and are not cost-effective. However, once signs of metastasis exist, favored sites of spread, such as the liver, lungs, bone, and brain, should be scanned. Appropriate evaluations place patients into four clinical stages (Table 86-3).

TREATMENT

Surgical Management For a newly diagnosed cutaneous melanoma, wide surgical excision of the lesion with a margin of normal skin is necessary to remove all malignant cells and minimize local recurrence. The appropriate width of the margin is a source of controversy. Based upon clinical studies, the following margins can be recommended for primary melanoma: in situ: 0.5 cm; invasive up to 1 mm thick: 1.0 cm; 1 to 4 mm thick: 2.0 cm; >4 mm thick: at least 2 cm. For lesions on the face, hands, and feet, strict adherence to these margins must give way to individual considerations about the constraints of surgery and minimization of morbidity. In all instances, however, inclusion of subcutaneous fat in the surgical specimen facilitates adequate thickness measurement and assessment of surgical margins by the pathologist.

Elective Regional Node Dissection Elective regional node dissection inAJCCstage II disease (without palpable adenopathy) has been advocated, based on the hypothesis that melanoma metastasizes in an orderly fashion from the skin to regional lymph nodes and finally to distant sites. If that is the case, surgical excision of nodal micrometastases could theoretically provide definitive treatment at a time of relatively low tumor burden and perhaps improve survival. The efficacy of this procedure remains controversial; while some retrospective series suggest a survival benefit, randomized studies examining this question showed no survival advantage for wide local excision followed by immediate elective regional node dissection compared with wide local excision followed by delayed dissection (only if nodes became palpable). Furthermore, the procedure has associated morbidity, and some lesions, especially those on the trunk, have ambiguous nodal draining sites, making it difficult to decide which area to dissect.

Results of biopsy of the first drainage node -- the so-called sentinel node -- predicts the likelihood of metastases in higher nodes. Sentinel nodes can be identified by injecting a blue dye or radioactive isotope around the primary tumor site. A negative biopsy result appears to obviate the need for elective regional nodal dissection. Patients with lesions <1 mm thick have an excellent prognosis and need no node dissection; at the other extreme, patients with lesions >4 mm thick have such a high risk for distant metastases that elective node dissection may not alter the ultimate clinical outcome. A subset of patients with AJCC stage II lesions of intermediate thickness may benefit from elective regional node dissection, but there is no consensus about which patients should undergo this procedure.

Adjuvant Therapy For patients who are free of disease but at high risk for metastases, adjuvant therapy that complements surgery is needed to destroy occult micrometastases, prolong disease-free survival, and improve the cure rate. Many strategies have been tried unsuccessfully. However, adjuvant interferon-a may improve disease-free and overall survival, particularly in patients with nodal metastases (stage III disease). High-dose interferon, 20 million units per square meter intravenously 5 days a week for 4 weeks followed by 10 million units per square meter subcutaneously three times a week for 11 months, has been effective in some, but not all, studies. In nearly half of patients, these doses of interferon are associated with severe toxicity, including flulike illness and decline in performance status. The toxicity reverses with lower doses and when therapy is stopped. If interferon is beneficial at all, it benefits only a small fraction of treated patients.

Treatment of Metastatic Disease Melanoma can metastasize to any organ, the brain being a particularly common site. Metastatic melanoma generally is incurable, with survival in patients with visceral metastases generally<1 year. Thus, the goal of treatment is usually palliation. Patients with soft-tissue and node metastases fare better than those with liver and brain metastases. Metastases limited to regional nodes (AJCCstage III disease) warrant a therapeutic lymph node dissection. Surgical excision of a single metastasis to the lung or to a surgically accessible brain site can prolong survival. Trials of stereotactic radiosurgery will determine its future role in the treatment of brain metastases. More often, however, patients have multiple brain metastases that require radiation therapy and glucocorticoids. Radiation therapy can provide local palliation for recurrent tumors or metastases. Patients who have advanced regional disease limited to a limb may benefit from hyperthermic limb perfusion with melphalan and tumor necrosis factor. Complete response rates >90% have been reported; responses are associated with significant palliation of symptoms.

A number of drugs and biologicals have minimal antitumor activity (15 to 20% partial response rates) in metastatic melanoma, including dacarbazine (DTIC); the nitrosoureas carmustine (BCNU), lomustine (CCNU), and semustine (methyl-CCNU); platinum analogues such as cisplatin and carboplatin; vinca alkyloids such as vincristine, vinblastine, and vindesine; the taxanes paclitaxel and docetaxel; interferon; and interleukin 2 (IL-2). Single-agent dacarbazine is considered the standard treatment. This agent has been given at a number of different doses and schedules; 250 mg/m$_2$intravenously every day for 5 days every 3 weeks is a standard schedule. Dacarbazine-based combination regimens are probably more effective. Interferon and IL-2 produce response rates similar to those seen with cytotoxic agents; however, at

active doses, they usually cause greater toxicity than chemotherapy.

Melanomas often express cell-surface antigens that may be recognized by host immune cells. A number of melanoma-associated antigens have been discovered. Melanoma antigens (MAGEs)-1, -2, and -3 (endogenous proteins controlled by genes on the X chromosome; there are up to 12 of these genes) and tyrosinase, an enzyme involved in melanin synthesis, are antigens that are processed into peptides and presented to T cells via HLA-A antigens on the tumor, particularly the HLA-A1 and -A2 alleles, which are expressed in about 85% of patients with melanoma. In addition, a melanoma antigen called MART is recognized in the context of class II MHC antigens. These melanoma-associated antigens alone or in combination may make it possible to develop vaccination strategies against melanoma. Such strategies include the use of purified proteins as immunogens and the use of genetically altered tumor cells to elicit a T cell response. Alternative experimental approaches include efforts to expand tumor-specific T cells (obtained either from the tumor as tumor-infiltrating lymphocytes or harvested from the peripheral blood after vaccination) in vitro and transfer them into patients in large numbers. In addition, monoclonal antibodies to tumor antigens are being tested, with some early indication of efficacy in around 15% of patients. All of these experimental approaches will need considerable further development before being applicable on a wide scale. However, once an approach is found that is active in metastatic disease, it may prove most useful as adjuvant therapy.

The absence of curative therapy for patients with metastatic melanoma underscores the importance of early detection and prevention as strategies to decrease melanoma mortality.

NONMELANOMA SKIN CANCER

Nonmelanoma skin cancer is the most common cancer in the United States, with an estimated annual incidence of more than 1,000,000 cases. Basal cell carcinomas (BCCs) account for 70 to 80% of nonmelanoma skin cancers. Squamous cell carcinomas (SCCs), while representing only about 20% of nonmelanoma skin cancers, are more significant because of their ability to metastasize; they account for most of the 2300 deaths annually. Incidence rates have risen dramatically over the past decade.

ETIOLOGY

The cause ofBCC andSCC is multifactorial. Cumulative exposure to sunlight, principally the ultraviolet B (UV-B) spectrum, is the most significant factor. Other factors associated with a higher incidence of skin cancer are male sex, older age, Celtic descent, a fair complexion, a tendency to sunburn easily, and an outdoor occupation. The incidence of these tumors increases with decreasing latitude. Most tumors develop on sun-exposed areas of the head and neck. Tumors are more common on the left side of the body in the United States but on the right side in England, presumably owing to asymmetric exposure during driving. As the earth's protective ozone shield continues to thin, further increases in the incidence of skin cancer can be anticipated. In certain geographic areas, exposure to arsenic in well water or from industrial sources may significantly increase the risk of BCC and SCC. Skin cancer in affected individuals may be seen with or without other cutaneous markers of chronic arsenism (e.g., arsenical keratoses).

Less common is exposure to the cyclic aromatic hydrocarbons in tar, soot, or shale. The risk of lip or oral SCC is increased with cigarette smoking. Human papillomaviruses and ultraviolet radiation may act as cocarcinogens.

Host factors associated with a high risk of skin cancer include immunosuppression induced by disease or drugs. Transplant recipients receiving chronic immunosuppressive therapy are particularly prone toSCC. The frequency of skin cancer is proportional to the duration of immunosuppression and the extent of sun exposure. Skin cancer is a not uncommon finding in patients infected with HIV, and it may be more aggressive in this setting. Other factors include ionizing radiation, thermal burns, certain scars, and chronic ulcerations. Several heritable conditions have been associated with skin cancer (e.g., albinism, xeroderma pigmentosum, andBCCnevus syndrome). Mutations in the tumor suppressor gene, *patched*, may lead to BCC.

CLINICAL PRESENTATION

Nonmelanoma skin cancers are often asymptomatic, but nonhealing ulceration, bleeding, or pain can occur.

Basal Cell Carcinoma BCC is a malignancy arising from epidermal basal cells. The most common type is *noduloulcerative BCC*, which begins as a small, pearly nodule, often with small telangiectatic vessels on its surface. The nodule grows slowly and may undergo central ulceration. Various amounts of melanin may be present in the tumor; tumors with a heavier accumulation are referred to as *pigmented BCC*. While clinically no more aggressive than the noduloulcerative variant, the latter may be mistaken for malignant melanoma. *Superficial BCC* consists of one or several erythematous, scaling plaques that slowly enlarge. Although they are more commonly found on the trunk and extremities, the head and neck can also be affected. The lesions may be confused with benign inflammatory dermatoses, especially nummular eczema and psoriasis. *Morpheaform (fibrosing) BCC* manifests itself as a solitary, flat or slightly depressed, indurated, whitish or yellowish plaque. Borders are typically indistinct, a feature associated with a greater potential for extensive subclinical spread.

Squamous Cell Carcinoma Primary cutaneousSCC is a malignant neoplasm of keratinizing epidermal cells. UnlikeBCC, which has a very low metastatic potential, SCC can metastasize and grow rapidly. The clinical features of SCC vary widely. Commonly, SCC appears as an ulcerated nodule or a superficial erosion on the skin or lower lip, but it may present as a verrucous papule or plaque. Unlike BCC, overlying telangiectasias are uncommon. The margins of this tumor may be ill-defined, and fixation to underlying structures may occur. Cutaneous SCC may develop anywhere on the body, but it usually arises on sun-damaged skin. A related neoplasm, keratoacanthoma, typically appears as a dome-shaped papule with a central keratotic crater, expands rapidly, and commonly regresses without therapy. This lesion can be difficult to differentiate from SCC.

SCChas several premalignant forms (actinic keratosis, actinic cheilitis, and some cutaneous horns) and in situ forms (e.g., Bowen's disease) that are confined to the epidermis. Actinic keratoses and cheilitis are hyperkeratotic papules and plaques that occur on sun-exposed areas. While the potential for malignant degeneration is low in

any individual lesion, the risk of SCC increases with larger numbers of lesions. Bowen's disease presents as a scaling, erythematous plaque, which may develop into invasive SCC in up to 20% of cases. Controversy exists regarding the association of Bowen's disease with internal malignancy; however, no significant relationship is noted when other predisposing factors (e.g., arsenic) are absent. Treatment of premalignant and in situ lesions reduces the subsequent risk of invasive disease.

NATURAL HISTORY

Basal Cell Carcinoma The natural history ofBCC is that of a slowly enlarging, locally invasive neoplasm. The degree of local destruction and risk of recurrence vary with the size, duration, and location of the tumor; the histologic subtype; the presence of recurrent disease; and various patient characteristics. Location on the central face (e.g., the nose, the nasolabial fold, or the periorbital or perioral area), the ears, or the scalp may portend a higher risk. Small nodular, pigmented, cystic, or superficial BCCs respond well to most treatments. Large nodular, noduloulcerative, and especially morpheaform BCCs may be more aggressive. The metastatic potential of BCC is about 0.0028 to 0.1%. Persons with either BCC orSCC have an increased risk of developing subsequent skin cancers.

Squamous Cell Carcinoma The natural history ofSCCdepends on both tumor and host characteristics. Tumors arising on actinically damaged skin have a lower metastatic potential than those on protected surfaces. The metastatic frequency of cutaneous SCC, 0.3 to 3.7%, is lower than that of mucosal SCC. Tumors occurring on the lower lip and ear have metastatic potential approaching 13 and 11%, respectively. The metastatic potential of SCC arising in burn scars, chronic ulcerations, or the genitalia is higher. The overall metastatic rate for recurrent tumors approaches 30%. Poorly differentiated, deep tumors with perineural or lymphatic invasion often behave aggressively. Multiple tumors with rapid growth and aggressive behavior can be a therapeutic challenge in immunosuppressed patients. Regional lymph nodes are the most common site of metastasis. In patients with metastatic disease, the 5-year survival rate may be low.

TREATMENT

Basal Cell Carcinoma The treatment modalities used forBCCinclude electrodesiccation and curettage (ED&C), excision, cryosurgery, radiation therapy, Mohs micrographic surgery (MMS), and others. The mode of therapy chosen depends on tumor characteristics, age, medical status, preferences of the patient, and other factors. ED&C remains the method most commonly employed by dermatologists. This method is selected for low-risk tumors (e.g., a small primary tumor of a less aggressive subtype in a favorable location). Excision, which offers the advantage of histologic control, is usually selected for more aggressive tumors or those in high-risk locations, or, in many instances, for esthetic reasons. Cryosurgery using liquid nitrogen may be used in certain low-risk tumors, but it requires specialized equipment (cryoprobes) to be effective for advanced neoplasms. Radiation therapy, while not employed as often as surgical modalities, offers an excellent chance for cure in many cases of BCC. It is useful in patients not considered surgical candidates and as a surgical adjunct in high-risk tumors. Younger patients may not be good candidates for radiation therapy because of

the risks of long-term carcinogenesis and radiodermatitis. MMS is a specialized type of surgical excision that permits the ultimate in histologic control and preservation of uninvolved tissue. It is preferred for lesions that are recurrent, in a high-risk location, or large and ill-defined, and where maximal tissue conservation is critical (e.g., the eyelids). Topical chemotherapy with 5-fluorouracil (5FU) cream has limited usefulness in the management of BCC and should be used only for treating superficial BCC. Intralesional 5FU is being investigated for BCC. Intralesional interferon is effective in certain primary tumors. Photodynamic therapy, which employs selective activation of a photoactive drug by visible light, may be useful in patients with numerous tumors. Lasers also have been used for the treatment of skin cancer.

Squamous Cell Carcinoma The therapy of cutaneousSCCshould be based on an analysis of risk factors influencing the biologic behavior of the tumor. These include the size, location, and degree of histologic differentiation of the tumor and the age and physical condition of the patient. Surgical excision,MMS, and radiation are standard methods of treatment. Cryosurgery andED&Chave been used successfully for small primary tumors. Metastases are treated with lymph node dissection, irradiation, or both. 13-*Cis*-retinoic acid (1 mg orally every day) plus interferon (3 million units subcutaneously or intramuscularly every day) may produce a partial response in most patients. Systemic chemotherapy combinations that include cisplatin may also be palliative in some patients.

PREVENTION

Since the vast majority of skin cancers are related to chronicUV-Bexposure, they are largely preventable by blocking sun exposure. Emphasis should be placed on preventive measures beginning early in life. Patients must understand that damage from UV-B begins early, despite the fact that cancers develop years later. Regular use of sunscreens and protective clothing should be encouraged. Avoidance of tanning salons and sun exposure during midday (10 A.M. to 2 P.M.) is recommended. Precancerous and in situ lesions should be treated early. Early detection of small tumors affords simpler treatment modalities with higher cure rates and lower morbidity. In patients with a history of skin cancer, long-term follow-up for the detection of recurrence, metastasis, and new skin cancers should be emphasized. Chemoprophylaxis using synthetic retinoids is useful in controlling new lesions in some patients with multiple tumors.

OTHER TYPES OF CUTANEOUS CANCER

Neoplasms of cutaneous adnexa, and sarcomas of fibrous, mesenchymal, fatty, and vascular tissues make up 1 to 2% of nonmelanoma skin cancers. The recent rapid rise in the incidence of Kaposi's sarcoma is attributed to HIV infection and immunosuppressive therapy. Human herpesvirus 8 appears to be the cause of sporadic and HIV-associated Kaposi's sarcoma. Current therapy is palliative and depends on the symptoms and sites of involvement. Treatment modalities include cryosurgery, vinblastine, excision, radiation, interferon, and systemic combination chemotherapy (Chap. 309).

ACKNOWLEDGEMENT
The authors wish to acknowledge Dr. Nhu-linh T. Tran, who was a co-author of

this chapter in the 14th edition.

(Bibliography omitted in Palm version)

Back to Table of Contents

87. HEAD AND NECK CANCER - *Everett E. Vokes*

Epithelial carcinomas of the head and neck arise from the mucosal surfaces in the head and neck area and typically are squamous cell in origin. This category includes tumors of the paranasal sinuses, the oral cavity, and the nasopharynx, oropharynx, hypopharynx, and larynx. Tumors of the salivary glands differ from the more common carcinomas of the head and neck in etiology, histopathology, clinical presentation, and therapy.**Thyroid malignancies are described in Chap. 330.*

INCIDENCE AND EPIDEMIOLOGY

The annual number of new cases of head and neck cancers in the United States is approximately 40,000, accounting for about 5% of adult malignancies. Head and neck cancers are more common in certain other countries, and the worldwide incidence exceeds half a million cases annually. In North America and Europe, the tumors usually arise from the oral cavity, oropharynx, or larynx, whereas nasopharyngeal cancer is more common in the Mediterranean countries and in the Far East.

ETIOLOGY AND GENETICS

Alcohol and tobacco use are the most common risk factors for head and neck cancer in the United States. Smokeless tobacco is an etiologic agent for oral cancers. Other potential carcinogens include marijuana and occupational exposures such as nickel refining, exposure to textile fibers, and woodworking.

Dietary factors may contribute. The incidence of head and neck cancer is highest in people with the lowest consumption of fruits and vegetables. Certain vitamins, including dietary carotenoids, may be protective; retinoids are being tested for prevention.

Some head and neck cancers may have a viral etiology. The DNA of human papilloma virus has been detected in the tissue of oral and tonsil cancers, and Epstein-Barr virus (EBV) infection is associated with nasopharyngeal cancer. Nasopharyngeal cancer occurs endemically in some countries of the Mediterranean and Far East, where EBV antibody titers can be measured to screen high-risk populations. Nasopharyngeal cancer has also been associated with other environmental factors, such as consumption of salted fish.

No specific risk factors or environmental carcinogens have been identified for salivary gland tumors.

HISTOPATHOLOGY, CARCINOGENESIS, AND MOLECULAR BIOLOGY

Squamous cell head and neck carcinomas can be divided into well-differentiated, moderately well-differentiated, and poorly differentiated categories. Patients with poorly differentiated tumors have a worse prognosis than those with well-differentiated tumors. For nasopharyngeal cancers, the less common differentiated squamous cell carcinoma is distinguished from nonkeratinizing and undifferentiated carcinoma (lymphoepithelioma) that contains infiltrating (bystander) lymphocytes.

Salivary gland tumors can arise from the major (parotid, submandibular, sublingual) or minor salivary glands (located in the submucosa of the upper aerodigestive tract). Most parotid tumors are benign, but half of submandibular and sublingual gland tumors and most minor salivary gland tumors are malignant. Malignant tumors include mucoepidermoid and adenoidcystic carcinomas and adenocarcinomas.

The mucosal surface of the entire pharynx is exposed to alcohol and tobacco-related carcinogens and is at risk for the development of a premalignant or malignant lesion, such as erythroplakia or leukoplakia (hyperplasia, dysplasia), that can progress to invasive carcinoma. Alternatively, multiple synchronous or metachronous cancers can develop. In fact, patients with early-stage head and neck cancer are at greater risk of dying of a second malignancy than of dying from a recurrence of the primary disease.

Second head and neck malignancies are not therapy-induced but, instead, reflect the exposure of the upper aerodigestive mucosa to the same carcinogens that caused the first cancer. These second primaries develop in the head and neck area, the lung, or the esophagus.

Chromosomal deletions and other alterations, most frequently involving chromosomes 3p, 9p, 17p, and 13q, have been identified in both premalignant and malignant head and neck lesions, as have mutations in tumor suppressor genes, commonly the *p53* gene. Amplification of oncogenes is less common, but overexpression of PRAD-1/bcl-1 (cyclin D1), bc1-2, transforming growth factor b, and the epidermal growth factor receptor have been described. The latter finding correlates positively with tumor size and poor outcome and is a target for experimental treatments.

Resected tumor specimens with histopathologically negative margins ("complete resection") can have histopathologically undetectable residual tumor cells with persistent *p53* mutations at the margins. Thus, a tumor-specific *p53* mutation can be detected in some phenotypically "normal" surgical margins, indicating residual disease. Patients with such submicroscopic marginal involvement may have a worse prognosis than patients with negative margins.

CLINICAL PRESENTATION AND DIFFERENTIAL DIAGNOSIS

Most head and neck cancers occur after age 50, although these cancers can appear in younger patients, including those without known risk factors. The manifestations vary according to the stage and primary site of the tumor. Patients with nonspecific signs and symptoms in the head and neck area should be evaluated with a thorough otolaryngologic exam, particularly if symptoms persist longer than 2 to 4 weeks.

Cancer of the nasopharynx typically does not cause early symptoms. However, on occasion it may cause unilateral serous otitis media due to obstruction of the eustachian tube, unilateral or bilateral nasal obstruction, or epistaxis. Advanced nasopharyngeal carcinoma causes neuropathies of the cranial nerves.

Carcinomas of the oral cavity present as nonhealing ulcers, changes in the fit of dentures, or painful lesions. Tumors of the tongue base or oropharynx can cause decreased tongue mobility and alterations in speech. Cancers of the oropharynx or

hypopharynx rarely cause early symptoms, but they may cause sore throat and/or otalgia.

Hoarseness may be an early symptom of laryngeal cancer, and persistent hoarseness requires referral to an otorhinolaryngologist for indirect laryngoscopy and/or radiographic studies. If a head and neck lesion treated initially with antibiotics does not resolve in a short period, further workup is indicated; to simply continue the antibiotic treatment may be to lose the chance of early diagnosis of a malignancy.

Advanced head and neck cancers in any location can cause severe pain, otalgia, airway obstruction, cranial neuropathies, trismus, odynophagia, dysphagia, decreased tongue mobility, fistulas, skin involvement, and massive cervical lymphadenopathy, which may be unilateral or bilateral. Some patients have enlarged lymph nodes even though no primary lesion can be detected by endoscopy or biopsy; these patients are considered to have carcinoma of unknown primary. If the enlarged nodes are located in the upper neck and the tumor cells are of squamous cell histology, the malignancy probably arose from a mucosal surface in the head or neck. Tumor cells in supraclavicular lymph nodes may also arise from a primary site in the chest or abdomen.

The physical examination should include scrutiny of all visible mucosal surfaces and palpation of the floor of mouth and tongue and of the neck. In addition to tumors themselves, leukoplakia -- a white mucosal patch -- or erythroplakia -- a red mucosal patch -- may be observed; these "premalignant" lesions can represent hyperplasia, dysplasia, or carcinoma in situ. All visible lesions should be biopsied. Further examination should be performed by the otorhinolaryngologist. Additional staging procedures include computed tomography of the head and neck to identify the extent of the disease. Patients with lymph node involvement should have chest radiography and a bone scan to screen for distant metastases. The definitive staging procedure is an endoscopic examination under anesthesia, which may include laryngoscopy, esophagoscopy, and bronchoscopy; during this procedure, multiple biopsy samples are obtained to establish a primary diagnosis, define the extent of primary disease, and identify any additional premalignant lesions or second primaries.

Head and neck tumors are classified according to the TNM system of the American Joint Committee on Cancer. This classification varies according to the specific anatomic subsite (Tables 87-1 and87-2). Distant metastases are found in <10% of patients at initial diagnosis, but in autopsy series, microscopic involvement of the lungs, bones, or liver is more common, particularly in patients with advanced neck lymph node disease.

In patients with lymph node involvement and no visible primary, the diagnosis should be made by lymph node excision. If the results indicate squamous cell carcinoma, a panendoscopy should be performed, with biopsy of all suspicious-appearing areas and directed biopsies of common primary sites, such as the nasopharynx, tonsil, tongue base, and pyriform sinus.

TREATMENT

Generally, patients with head and neck cancer can be categorized into three clinical groups: those with localized disease, those with locally or regionally advanced disease,

and those with recurrent and/or metastatic disease. Comorbidities associated with tobacco and alcohol abuse can affect treatment outcome.

Localized Disease Approximately one-third of patients have localized disease; that is, T1 or T2 (stage I or stage II) lesions without detectable lymph node involvement or distant metastases. These lesions are treated with curative intent by surgery or radiation. The choice of modality differs according to institutional expertise. Generally, radiation therapy is preferred for laryngeal cancer to preserve voice function, and surgery is preferred for small lesions in the oral cavity to avoid the long-term complications of radiation, such as xerostomia and dental decay. Overall 5-year survival is 60 to 90%.

Locally or Regionally Advanced Disease Locally or regionally advanced disease -- that is, disease with a large primary tumor and/or lymph node metastases -- can also be treated with curative intent, but not with surgery or radiation therapy alone. Combined modality therapy including surgery, radiation therapy and chemotherapy is most successful. Concomitant chemotherapy and radiation therapy appears to be the most effective sequencing of treatment.

Induction Chemotherapy In this strategy, patients receive chemotherapy [usually cisplatin and fluorouracil (5FU)] before surgery and radiotherapy. Most patients who receive three cycles of this combination show tumor reduction, and the response is clinically "complete" in up to half of these patients. This "sequential" multimodality therapy does not cure more patients than surgery plus radiation therapy alone. Time to recurrence may be improved but survival is similar. However, induction chemotherapy allows for organ preservation in patients with laryngeal and hypopharyngeal cancer.

Concomitant Chemoradiotherapy With the concomitant strategy, chemotherapy and radiation therapy are given simultaneously rather than sequentially. Because most patients with head and neck cancer develop recurrent disease in the head and neck area, this approach is aimed at killing radiation-resistant cancer cells with chemotherapy. In addition, chemotherapy can enhance cell killing by radiation therapy. Toxicity (mucositis) is increased with concomitant chemoradiotherapy; however, meta-analysis of randomized trials documents an improvement in 5-year survival of 8% with concomitant5FU and radiation therapy. Results seem even better with 5FU and cisplatin plus radiation therapy. Five-year survival is 34 to 50%. The use of radiation therapy together with cisplatin has produced markedly improved survival in patients with advanced nasopharyngeal cancer. The success of concomitant chemoradiotherapy in patients with unresectable disease has led to the testing of a similar approach in patients with resectable disease in an effort to increase organ preservation.

Recurrent and/or Metastatic Disease Patients with recurrent and/or metastatic disease are, with few exceptions, treated with palliative intent. Some patients may require local or regional radiation therapy for pain control, but most are given chemotherapy. Response rates to chemotherapy average only 30 to 50%; the duration of response averages only 3 months, and the median survival time is 6 months. Therefore, chemotherapy provides transient symptomatic benefit. Drugs with single-agent activity in this setting include methotrexate,5FU, cisplatin, paclitaxel, and docetaxel. Combinations of cisplatin and 5FU, carboplatin and 5FU, and cisplatin and

paclitaxel are also used.

CHEMOPREVENTION

b-carotene and *cis*-retinoic acid can lead to the regression of leukoplakia. In addition, the use of *cis*-retinoic acid may reduce the incidence of second primaries.

TREATMENT COMPLICATIONS

Complications involved in the treatment of head and neck cancer are usually related to the extent of surgery. Several attempts have been made to limit the extent of surgery or to replace it with chemotherapy and radiation therapy. Acute complications of radiation include mucositis and dysphagia. Long-term complications include xerostomia, loss of taste, decreased tongue mobility, second malignancies, and dysphagia and neck fibrosis. The complications of chemotherapy vary with the regimen used but usually include myelosuppression, mucositis, nausea and vomiting, and nephrotoxicity (with cisplatin).

SALIVARY GLAND TUMORS

Most benign salivary gland tumors are treated with surgical excision, and patients with invasive salivary gland tumors are treated with surgery and radiation therapy. Neutron radiation may be particularly effective. These tumors may recur regionally; adenoidcystic carcinoma has a tendency to recur along the nerve tracks. Distant metastases may occur as late as 10 to 20 years after the initial diagnosis. For metastatic disease, therapy is given with palliative intent, usually chemotherapy with doxorubicin and/or cisplatin.

(Bibliography omitted in Palm version)

Back to Table of Contents

88. NEOPLASMS OF THE LUNG - *John D. Minna*

Each year, primary carcinoma of the lung affects 94,000 males and 78,000 females in the United States, 86% of whom die within 5 years of diagnosis, making it the leading cause of cancer death in both men and women and in all races. The incidence of lung cancer peaks between ages 55 and 65 years. Lung cancer accounts for 31% of all cancer deaths in men and 25% in women. The effects of smoking cessation efforts begun 25 years ago have been seen in a slowing of the rate of age-adjusted cancer death from lung cancer in males (~70 per 100,000 male population); but, unfortunately, the rate in females is still increasing (~35 per 100,000 female population). Only 15% of patients have local disease at diagnosis; 25% have disease spread to regional lymph nodes, and>55% have distant metastases. The 5-year survival rate of patients with local disease is 50%; it is 20% for patients with regional disease and 14% overall. The 5-year overall lung cancer survival rate has nearly doubled in the last 30 years. The improvement is due to advances in combined-modality treatment with surgery, radiotherapy, and chemotherapy. Thus, primary carcinoma of the lung is a major health problem with a generally grim prognosis. However, an orderly approach to diagnosis, staging, and treatment based on knowledge of the clinical behavior of lung cancer and involving multidisciplinary input allows choice and delivery of the best therapy for potential cure or optimal palliation of individual patients.

PATHOLOGY

The term *lung cancer* is used for tumors arising from the respiratory epithelium (bronchi, bronchioles, and alveoli). Mesotheliomas, lymphomas, and stromal tumors (sarcomas) are distinct from epithelial lung cancer. Four major cell types make up 88% of all primary lung neoplasms according to the World Health Organization classification (Table 88-1). These are *squamous* or *epidermoid carcinoma*, *small cell* (also called *oat cell*) *carcinoma*, *adenocarcinoma* (including bronchioloalveolar), and *large cell* (also called *large cell anaplastic*) *carcinoma*. The remainder include undifferentiated carcinomas, carcinoids, bronchial gland tumors (including adenoid cystic carcinomas and mucoepidermoid tumors), and rarer tumor types. The various cell types have different natural histories and responses to therapy, and thus a correct histologic diagnosis by an experienced pathologist is the first step to correct treatment. In the past 25 years, for unknown reasons, adenocarcinoma has replaced squamous cell carcinoma as the most frequent histologic subtype for all sexes and races combined (Table 88-1).

Major treatment decisions are made on the basis of whether a tumor is classified histologically as a small cell carcinoma or as one of the non-small cell varieties (epidermoid, adenocarcinoma, large cell carcinoma, bronchioloalveolar carcinoma, and mixed versions of these). Some of the distinctions are summarized in Tables 88-1 and88-2. At presentation, small cell carcinomas usually have already spread such that surgery is unlikely to be curative, and they are managed primarily by chemotherapy with or without radiotherapy. In contrast, non-small cell cancers that are found to be localized at the time of presentation may be cured with either surgery or radiotherapy. Non-small cell cancers do not respond as well to chemotherapy as small cell cancers.

Ninety percent of patients with lung cancer of all histologic types are current or former cigarette smokers. Of the annual 171,600 new cases of lung cancer, ~50% develop in

former smokers. With increased success in smoking cessation efforts, the number of former smokers will grow, and these individuals will be important candidates for early detection and chemoprevention efforts. By far the most common form of lung cancer arising in lifetime nonsmokers, in women, and in young patients (<45 years) is adenocarcinoma. However, in nonsmokers with adenocarcinoma involving the lung, the possibility of other primary sites should be considered. Epidermoid and small cell cancers usually present as central masses with endobronchial growth, while adenocarcinomas and large cell cancers tend to present as peripheral nodules or masses, frequently with pleural involvement. Epidermoid and large cell cancers cavitate in ~10 to 20% of cases. Bronchioloalveolar carcinoma, a form of adenocarcinoma arising from peripheral airways, can present as a single mass; as a diffuse, multinodular lesion; or as a fluffy infiltrate.

ETIOLOGY

Most lung cancers are caused by carcinogens and tumor promoters ingested via cigarette smoking. The prevalence of smoking in the United States is 28% for males and 25% for females, age 18 years or older; 38% of high school seniors smoke. The relative risk of developing lung cancer is increased about 13-fold by active smoking and about 1.5-fold by long-term passive exposure to cigarette smoke. Chronic obstructive pulmonary disease, which is also smoking-related, further increases the risk of developing lung cancer. The lung cancer death rate is related to the total amount (often expressed in "cigarette pack-years") of cigarettes smoked, such that the risk is increased 60- to 70-fold for a man smoking two packs a day for 20 years as compared with a nonsmoker. Conversely, the chance of developing lung cancer decreases with cessation of smoking but may never return to the nonsmoker level. The increase in lung cancer rate in women is also associated with a rise in cigarette smoking. Women have a higher relative risk per given exposure than men (~1.5 fold higher), and women with lung cancer are more likely to have never smoked than men. This gender difference is likely due to a higher susceptibility to tobacco carcinogens in women.

Efforts to get people to stop smoking are mandatory. However, smoking cessation is extremely difficult, because the smoking habit represents a powerful addiction to nicotine (Chap. 390). Preventing people from starting to smoke may be more effective, an effort that needs to be targeted to children.

Although human lung cancer is not thought of as a genetic disease, various molecular genetic studies have shown the acquisition by lung cancer cells of a number of genetic lesions, including activation of dominant oncogenes and inactivation of tumor suppressor or recessive oncogenes (Chaps. 81 and82). In fact, lung cancer cells may have to accumulate a large number (perhaps³10) of such lesions. For the dominant oncogenes, these include point mutations in the coding regions of the *ras* family of oncogenes (particularly in the K-*ras* gene in adenocarcinoma of the lung); amplification, rearrangement, and/or loss of transcriptional control of *myc* family oncogenes (c-, N-, and L-*myc*; changes in c-*myc* are found in non-small cell cancers, while changes in all *myc* family members are found in small cell lung cancer); and overexpression of *bcl-2*, *Her-2/neu*, and the telomerase gene (Table 88-2). Tumor mutations in *ras* genes are associated with a poor prognosis in non-small cell lung cancer, while tumor amplification of c-*myc* is associated with a poor prognosis in small cell lung cancer.

For the recessive oncogenes (*tumor suppressor genes*), cytogenetic and allelotyping analyses have shown allele loss involving chromosome regions 1p, 1q, 3p12-13, 3p14 (*FHIT* gene region), 3p21, 3p24-25, 4p, 4q, 5q, 8p, 9p (*p16/CDKN2*, *p15,p19ARF* gene cluster), 11p13, 11p15, 13q14 (retinoblastoma, *rb*, gene), 16q, and 17p13 (*p53* gene), as well as other sites. Several candidate recessive oncogenes on chromosome 3p appear to be involved in nearly all lung cancers and may be affected early in preneoplastic lesions. The *p53* and *rb* genes are both mutated in>90% of small cell lung cancers, while *p53* is mutated in >50% and *rb* in 20% of non-small cell lung cancers. *p16/CDKN2* is abnormal in 10% of small cell and >50% of non-small cell lung cancers. Rb and *p16/CDKN2* are part of the same G1-to-S cell cycle regulatory pathway. Either one or the other of these elements appears to be mutated or to have its expression turned off (e.g., by hypermethylation of the promoter) in the large majority of lung cancers. Histologically identifiable preneoplastic lesions found in the respiratory epithelium of lung cancer patients and smokers include hyperplasia, dysplasia (progressively severe), and carcinoma in situ. 3p allele loss (hyperplasia) followed by 9p (*p16/CDKN2*) allele loss (hyperplasia) are the earliest events; 17p (*p53*) abnormalities and then *ras* mutations usually are found only in carcinoma in situ and invasive cancer. Thus, molecular changes involving allele loss and microsatellite alteration can be found in the earliest preneoplastic lesions and potentially even before any histologic changes are noted. Clinical trials of early diagnosis are needed to prove the usefulness of these molecular markers in the identification of very early lung cancer and in the monitoring of treatment and chemoprevention.

The large number of lesions shows that lung cancer, like other common epithelial malignancies, is a multistep process that is likely to involve both carcinogens and tumor promoters. Prevention can be directed at both processes. Lung cancer cells produce many peptide hormones and express receptors for these hormones, which can act to stimulate tumor cell growth in an "autocrine" fashion. Highly carcinogenic derivatives of nicotine are formed in cigarette smoke. Lung cancer cells of all histologic types express receptors for nicotine. Nicotine can prevent apoptosis in lung cancer cell lines. Thus, nicotine itself could be directly involved in lung cancer pathogenesis.

While lung cancer does not have a clear pattern of Mendelian inheritance, several features suggest a potential for familial association. Inherited mutations in *rb* (patients with retinoblastomas living to adulthood) and *p53* (Li-Fraumeni syndrome) genes may develop lung cancer. First-degree relatives of lung cancer probands have a two- to threefold excess risk of lung cancer or other cancers, many of which are not smoking-related. Genetic epidemiologic studies have proposed an association between the P450 enzyme or chromosome fragility (*mutagen sensitivity*) genotypes and the development of lung cancer. The identification of persons at very high risk of developing lung cancer would be useful in early detection and prevention efforts.

CLINICAL MANIFESTATIONS

Lung cancer gives rise to signs and symptoms caused by local tumor growth, invasion or obstruction of adjacent structures, growth in regional nodes through lymphatic spread, growth in distant metastatic sites after hematogenous dissemination, and remote effects of tumor products (paraneoplastic syndrome). Peptide hormone secretion

by the tumor or immunologic cross-reaction between tumor and normal tissue antigens can produce a variety of signs and symptoms (Chap. 100).

Although 5 to 15% of patients with lung cancer are identified while they are asymptomatic, usually as a result of a routine chest radiograph, most patients present with some sign or symptom. Central or endobronchial growth of the primary tumor may cause cough, hemoptysis, wheeze and stridor, dyspnea, and postobstructive pneumonitis (fever and productive cough). Peripheral growth of the primary tumor may cause pain from pleural or chest wall involvement, cough, dyspnea on a restrictive basis, and symptoms of lung abscess resulting from tumor cavitation. Regional spread of tumor in the thorax (by contiguous growth or by metastasis to regional lymph nodes) may cause tracheal obstruction, esophageal compression with dysphagia, recurrent laryngeal nerve paralysis with hoarseness, phrenic nerve paralysis with elevation of the hemidiaphragm and dyspnea, and sympathetic nerve paralysis with Horner's syndrome (enophthalmos, ptosis, miosis, and ipsilateral loss of sweating). *Pancoast's* (or *superior sulcus tumor*) *syndrome* results from local extension of a tumor (usually epidermoid) growing in the apex of the lung with involvement of the eighth cervical and first and second thoracic nerves, with shoulder pain that characteristically radiates in the ulnar distribution of the arm, often with radiologic destruction of the first and second ribs. Often Horner's syndrome and Pancoast's syndrome coexist. Other problems of regional spread include *superior vena cava syndrome* from vascular obstruction; pericardial and cardiac extension with resultant tamponade, arrhythmia, or cardiac failure; lymphatic obstruction with resultant pleural effusion; and lymphangitic spread through the lungs with hypoxemia and dyspnea. In addition, bronchioloalveolar carcinoma can spread transbronchially, producing tumor growing along multiple alveolar surfaces with impairment of gas exchange, respiratory insufficiency, dyspnea, hypoxemia, and sputum production.

Extrathoracic metastatic disease is found at autopsy in >50% of patients with epidermoid carcinoma, 80% of patients with adenocarcinoma and large cell carcinoma, and >95% of patients with small cell cancer. Lung cancer metastases may occur in virtually every organ system. Common clinical problems related to metastatic lung cancer include brain metastases with neurologic deficits; bone metastases with pain and pathologic fractures; bone marrow invasion with cytopenias or leukoerythroblastosis; liver metastases causing biochemical liver dysfunction, biliary obstruction, and pain; lymph node metastases in the supraclavicular region and occasionally in the axilla and groin; and spinal cord compression syndromes from epidural or bone metastases. Adrenal metastases are common but rarely cause adrenal insufficiency.

Paraneoplastic syndromes are common in patients with lung cancer and may be the presenting finding or first sign of recurrence. In addition, paraneoplastic syndromes may mimic metastatic disease and, unless detected, lead to inappropriate palliative rather than curative treatment. Often the paraneoplastic syndrome may be relieved with successful treatment of the tumor. In some cases, the pathophysiology of the paraneoplastic syndrome is known, particularly when a hormone with biologic activity is secreted by a tumor (Chap. 100). However, in many cases the pathophysiology is unknown. Systemic symptoms of anorexia, cachexia, weight loss (seen in 30% of patients), fever, and suppressed immunity are paraneoplastic syndromes of unknown etiology. *Endocrine syndromes* are seen in 12% of patients; hypercalcemia and

hypophosphatemia resulting from the ectopic production by epidermoid tumors of parathyroid hormone (PTH) or PTH-related peptide production, hyponatremia with the syndrome of inappropriate secretion of antidiuretic hormone or possibly atrial natriuretic factor by small cell cancer, and ectopic secretion by small cell cancer of adrenocorticotropic hormone (ACTH). ACTH secretion usually results in additional electrolyte disturbances, especially hypokalemia, rather than the changes in body habitus that occur in Cushing's syndrome from a pituitary adenoma.

Skeletal-connective tissue syndromes include clubbing in 30% of cases (usually non-small cell carcinomas) and hypertrophic pulmonary osteoarthropathy in 1 to 10% of cases (usually adenocarcinomas) with periostitis and clubbing giving pain, tenderness, and swelling over the affected bones and a positive bone scan. *Neurologic-myopathic syndromes* are seen in only 1% of patients but are dramatic and include the myasthenic *Eaton-Lambert syndrome* and retinal blindness with small cell cancer, while peripheral neuropathies, subacute cerebellar degeneration, cortical degeneration, and polymyositis are seen with all lung cancer types. Many of these are caused by autoimmune responses such as the development of anti-voltage-gated calcium channel antibodies in the Eaton-Lambert syndrome (Chap. 101). Coagulation, thrombotic, or other hematologic manifestations occur in 1 to 8% of patients and include migratory venous thrombophlebitis (*Trousseau's syndrome*), nonbacterial thrombotic (marantic) endocarditis with arterial emboli, disseminated intravascular coagulation with hemorrhage, and anemia, granulocytosis, and leukoerythroblastosis. Cutaneous manifestations such as dermatomyositis and acanthosis nigricans are uncommon (£1%), as are the renal manifestations of nephrotic syndrome or glomerulonephritis (£1%).

DIAGNOSIS AND STAGING

EARLY DIAGNOSIS

The screening of asymptomatic persons at high risk (men older than 45 years who smoke³40 cigarettes per day) by means of sputum cytology and chest radiographs has not improved the survival rate. Although 90% of patients whose lung cancer is detected by screening are asymptomatic, no difference was found in the survival rates of the screened and nonscreened groups. Women have not been studied. The use of low dose spiral computed tomography (CT) lung scanning may be more sensitive, particularly for peripheral lesions. However, false positive rates are high (25% have abnormal tests, only 10% of which are cancers), and survival benefit for screening has not yet been shown (Chap. 80).

ESTABLISHING A TISSUE DIAGNOSIS OF LUNG CANCER

Once signs, symptoms, or screening studies suggest lung cancer, a tissue diagnosis must be established. Tumor tissue can be obtained by a bronchial or transbronchial biopsy during fiberoptic bronchoscopy; by node biopsy during mediastinoscopy; from the operative specimen at the time of definitive surgical resection; by percutaneous biopsy of an enlarged lymph node, soft tissue mass, lytic bone lesion, bone marrow, or pleural lesion; by fine-needle aspiration of thoracic or extrathoracic tumor masses usingCTguidance; or from an adequate cell block obtained from a malignant pleural

effusion. In most cases, the pathologist should be able to make a definite diagnosis of epithelial malignancy and make the crucial differentiation of small cell from non-small cell lung cancer.

STAGING PATIENTS WITH LUNG CANCER

Lung cancer staging consists of two parts: first, a determination of the location of tumor (anatomic staging) and, second, an assessment of a patient's ability to withstand various antitumor treatments (physiologic staging). In a patient with non-small cell lung cancer, *resectability* (whether the tumor can be entirely removed by a standard surgical procedure such as a lobectomy or pneumonectomy), which depends on the anatomic stage of the tumor, and *operability* (whether the patient can tolerate such a surgical procedure), which depends on the cardiopulmonary function of the patient, are determined.

Non-Small Cell Lung Cancer The TNM International Staging System should be used for cases of non-small cell lung cancer, particularly in preparing patients for curative attempts with surgery or radiotherapy (Table 88-3). The various T (tumor size), N (regional node involvement), and M (presence or absence of distant metastasis) factors are combined to form different stage groups. At presentation, approximately one-third of patients have disease localized enough for a curative attempt with surgery or radiotherapy (patients with stage I or II disease and some with stage IIIA disease), one-third have distant metastatic disease (stage IV disease), and one-third have local or regional disease that may or may not be amenable to a curative attempt (some patients with stage IIIA disease and others with stage IIIB disease) (see below). This staging system provides useful prognostic information.

Small Cell Lung Cancer A simple two-stage system is used. In this system, limited-stage disease (seen in about 30% of all patients with small cell lung cancer) is defined as disease confined to one hemithorax and regional lymph nodes (including mediastinal, contralateral hilar, and usually ipsilateral supraclavicular nodes), while extensive-stage disease (seen in about 70% of patients) is defined as disease exceeding those boundaries. Clinical studies such as physical examination, x-rays,CT and bone scans, and bone marrow examination are used in staging. In part, the definition of limited-stage disease relates to whether the known tumor can be encompassed within a tolerable radiation therapy port. Thus, contralateral supraclavicular nodes, recurrent laryngeal nerve involvement, and superior vena caval obstruction can all be part of limited-stage disease. However, cardiac tamponade, malignant pleural effusion, and bilateral pulmonary parenchymal involvement generally qualify disease as extensive-stage because the organs within a curative radiation therapy port cannot safely tolerate curative radiation doses.

GENERAL STAGING PROCEDURES (See Table 88-4)

All patients with lung cancer should have a complete history and physical examination, with evaluation of all other medical problems, determination of performance status and history of weight loss, and aCT scan of the chest and abdomen with contrast. Positron emission tomography (PET) scans are sensitive in detecting metastatic disease. While not done in every patient, fiberoptic bronchoscopy provides material for pathologic

examination, information on tumor size, location, degree of bronchial obstruction (i.e., assesses resectability), and recurrence.

Chest radiographs andCT scans are needed to evaluate tumor size and nodal involvement; old radiographs are useful for comparison. CT scans are of use in the preoperative staging of non-small cell lung cancer to detect mediastinal nodes and pleural extension and occult abdominal disease (e.g., of the liver and adrenal glands), as well as in the planning of curative radiation therapy to allow the design of fields to encompass all the known tumor while avoiding as much normal tissue as possible. However, mediastinal nodal involvement should be documented histologically if the findings will influence therapeutic decisions. Thus, sampling of lymph nodes via mediastinoscopy or thoracotomy to establish the presence or absence of N2 or N3 nodal involvement is crucial in considering a curative surgical approach for patients with non-small cell lung cancer with clinical stage I, II, or III disease. Likewise, unless the CT-detected abnormalities are unequivocal, histology of suspicious abdominal lesions should be confirmed by procedures such as fine-needle aspiration if the patient would otherwise be considered for curative treatment. In small cell lung cancer, CT scans are used in the planning of chest radiation treatment and in the assessment of the response to chemotherapy and radiation therapy. Surgery or radiotherapy can make interpretation of conventional chest x-rays difficult; after treatment, CT scans can provide good evidence of tumor recurrence.

If signs or symptoms suggest involvement by tumor, brainCT or bone scans are performed, as well as radiography of any suspicious bony lesions. Any accessible lesions suspicious for cancer should be biopsied if a histologic diagnosis would influence treatment.

In patients presenting with a mass lesion on chest x-ray orCT scan and no obvious contraindications to a curative approach after the initial evaluation, the mediastinum must be investigated. Approaches vary among centers and include performing chest CT scan and mediastinoscopy (for right-sided tumors) or lateral mediastinotomy (for left-sided lesions) on all patients and proceeding directly to thoracotomy for staging of the mediastinum. In patients presenting with disease that is confined to the chest but not resectable, and who thus are candidates for neoadjuvant chemotherapy plus surgery or for curative radiotherapy with or without chemotherapy, other tests are done as indicated to evaluate specific symptoms. In patients presenting with non-small cell cancer that is not curable, all the general staging procedures are done, plus fiberoptic bronchoscopy as indicated to evaluate hemoptysis, obstruction, or pneumonitis, as well as thoracentesis with cytologic examination (and chest tube drainage as indicated) if fluid is present. As a rule, a radiographic finding of an isolated lesion (such as an enlarged adrenal gland) should be confirmed as cancer by fine-needle aspiration before a curative attempt is rejected.

STAGING OF SMALL CELL LUNG CANCER

Pretreatment staging for patients with small cell lung cancer includes the initial general lung cancer evaluation with chest and abdominalCT scans (because of the high frequency of hepatic and adrenal involvement) as well as fiberoptic bronchoscopy with washings and biopsies to determine the tumor extent before therapy; brain CT scan

(10% of patients have metastases); bone marrow biopsy and aspiration (20 to 30% of patients have tumor in the bone marrow); and radionuclide scans (bone) if symptoms or other findings suggest disease involvement in these areas. Chest and abdominal CT scans are very useful to evaluate and follow tumor response to therapy, and chest CT scans are helpful in planning chest radiotherapy ports.

If signs or symptoms of spinal cord compression or leptomeningitis develop at any time in lung cancer patients with disease of any histologic type, a spinalCT scan or magnetic resonance imaging (MRI) scan and examination of the cerebrospinal fluid cytology are performed. If malignant cells are detected, radiation therapy to the site of compression and intrathecal chemotherapy (usually with methotrexate) are given. In addition, a brain CT or MRI scan is performed to search for brain metastases, which often are associated with spinal cord or leptomeningeal metastases.

DETERMINATION OF RESECTABILITY AND OPERABILITY

In patients with non-small cell lung cancer, the following are major contraindications to curative surgery or radiotherapy alone: extrathoracic metastases; superior vena cava syndrome; vocal cord and, in most cases, phrenic nerve paralysis; malignant pleural effusion; cardiac tamponade; tumor within 2 cm of the carina (not curable by surgery but potentially curable by radiotherapy); metastasis to the contralateral lung; bilateral endobronchial tumor (potentially curable by radiotherapy); metastasis to the supraclavicular lymph nodes; contralateral mediastinal node metastases (potentially curable by radiotherapy); and involvement of the main pulmonary artery. Most patients with small cell lung cancer have unresectable disease; however, if clinical findings suggest the potential for resection (most common with peripheral lesions), that option should be considered.

PHYSIOLOGIC STAGING

Patients with lung cancer often have cardiopulmonary and other problems related to chronic obstructive pulmonary disease as well as other medical problems. To improve their preoperative condition, correctable problems (e.g., anemia, electrolyte and fluid disorders, infections, and arrhythmias) should be addressed, smoking stopped, and appropriate chest therapy instituted. Since it is not always possible to predict whether a lobectomy or pneumonectomy will be required until the time of operation, a conservative approach is to restrict resectional surgery to patients who could potentially tolerate a pneumonectomy. In addition to nonambulatory performance status, a myocardial infarction within the past 3 months is a contraindication to thoracic surgery because 20% of patients will die of reinfarction, while an infarction in the past 6 months is a relative contraindication. Other major contraindications include uncontrolled major arrhythmias, a maximum breathing capacity <40% of the predicted value, an FEV_1(forced expiratory volume in 1 s) <1 L, CO_2retention (which is more serious than hypoxemia), and severe pulmonary hypertension. Recommending surgery when the FEV_1 is 1.1 to 2.4 L requires careful judgment, while an FEV_1>2.5 L usually permits a pneumonectomy. In patients with borderline pulmonary status or a question of pulmonary hypertension, split pulmonary function testing by ventilation-perfusion lung scans can define physiologic operability. The activity from quantitative scans is summed for each lung in the anterior and posterior views, and the ratio of the normal to total lung

activity is multiplied by the FEV_1. Pneumonectomy usually is physiologically tolerable if this predicted value is>1 L.

TREATMENT

The overall treatment approach to patients with lung cancer is shown inTable 88-5. Patients should be encouraged to stop smoking. Those who do fare better than those who continue to smoke.

Non-Small Cell Lung Cancer: Localized Disease

Surgery In patients with non-small cell lung cancer of stages IA, IB, IIA and IIB (Table 88-3) who can tolerate operation, the treatment of choice is pulmonary resection. In stage IIIA cases where the patient's age, cardiopulmonary function, and anatomy are favorable, resection also should be considered. If a complete resection is possible, the 5-year survival rate for N1 disease is about 50%, while it is about 20% for N2 disease. However, only 20% of cases of N2 disease are technically resectable, and most of these are discovered to be N2 only at thoracotomy. Surgery for N2 disease is the most controversial area in the surgical management of lung cancer. Patients with N2 disease can be divided into "minimal" disease (involvement of only one node with microscopic foci, usually discovered at thoracotomy or mediastinoscopy) and the more common "advanced," bulky disease, clinically obvious onCT scans and discovered preoperatively. Patients with contralateral or bilateral positive mediastinal (N3) nodes, extracapsular nodal involvement, or fixed nodes are not considered candidates for resection. Approaches that may make resection possible include chest wall resection for direct extension of tumor, tracheal sleeve pneumonectomy, and sleeve lobectomy for lesions near the carina. Neoadjuvant (preoperative) chemotherapy has response rates of 50 to 60% and causes unresectable disease to become resectable in many patients who respond (see below). Video-assisted thoracic surgery (VATS) via thoracoscopy is not usually used for curative lung cancer resection but may be useful for peripheral lesions in patients with poor lung function.

The extent of resection is a matter of surgical judgment based on findings at exploration. Conservative resection that encompasses all known tumor gives survival equal to that obtained with more extensive procedures. However, lobectomy is superior to wedge resection in reducing the rate of local recurrence. Thus, lobectomy is preferred to pneumonectomy and wedge resection. Wedge resection and segmentectomy (potentially byVATS) are reserved for patients with poor pulmonary reserve and small peripheral lesions. About 43% of all patients with lung cancer undergo thoracotomy. Of these, 76% have a definitive resection, 12% are explored only for disease extent, and 12% have a palliative procedure with known disease left behind. About 30% of patients treated with resection for cure survive for 5 years, and 15% survive for 10 years (Table 88-3). The 30-day hospital mortality rate after pulmonary resection is 3% for lobectomy and 6% for pneumonectomy. Thus, most patients thought to have a "curative" resection ultimately die of metastatic disease (usually within 5 years of surgery).

Management of occult and stage 0 carcinomas In the uncommon situation where malignant cells are identified in a sputum or bronchial washing specimen but the chest radiograph appears normal (TX tumor stage), the lesion must be localized. More than

90% can be localized by meticulous examination of the bronchial tree with a fiberoptic bronchoscope under general anesthesia and collection of a series of differential brushings and biopsies. Often, carcinoma in situ or multicentric lesions are found in these patients. Current recommendations are for the most conservative surgical resection, allowing removal of the cancer and conservation of lung parenchyma, even if the bronchial margins are positive for carcinoma in situ. The 5-year overall survival rate for these occult cancers is ~60%. Close follow-up of these patients is indicated because of the high incidence of second primary lung cancers (5% per patient per year). One approach to in situ or multicentric lesions uses systemically administered hematoporphyrin (which localizes to tumors and sensitizes them to light) followed by bronchoscopic phototherapy.

Solitary pulmonary nodule When a patient presents with an asymptomatic, solitary pulmonary nodule (defined as an x-ray density completely surrounded by normal aerated lung, with circumscribed margins, of any shape, usually 1 to 6 cm in greatest diameter), a decision to resect or follow the nodule must be made. Approximately 35% of all such lesions in adults are malignant, most being primary lung cancer, while <1% are malignant in nonsmokers under 35 years of age. A complete history, including a smoking history, physical examination, routine laboratory tests, chestCT scan, fiberoptic bronchoscopy, and old chest x-rays are obtained.PETscans are useful in detecting lung cancers >1.5 cm in diameter. If no diagnosis is immediately apparent, the following risk factors would all argue strongly in favor of proceeding with resection to establish a histologic diagnosis: a history of cigarette smoking; age 35 years or older; a relatively large lesion; lack of calcification; chest symptoms; associated atelectasis, pneumonitis, or adenopathy; and growth of the lesion revealed by comparison with old x-rays. At present, only two radiographic criteria are reliable predictors of the benign nature of a solitary pulmonary nodule: lack of growth over a period >2 years and certain characteristic patterns of calcification. Calcification alone does not exclude malignancy. However, a dense central nidus, multiple punctate foci, and "bull's eye" (granuloma) and "popcorn ball" (hamartoma) calcifications are all highly suggestive of a benign lesion.

When old x-rays are not available and the characteristic calcification patterns are absent, the following approach is reasonable: Nonsmoking patients younger than 35 years can be followed with serialCT every 3 months for 1 year and then yearly. If any significant growth is found, a histologic diagnosis is needed. For patients older than 35 years and all patients with a smoking history, a histologic diagnosis must be made. The sample for histologic diagnosis can be obtained either at the time of nodule resection or, if the patient is a poor operative risk, viaVATS or transthoracic fine-needle biopsy. Some institutions use preoperative fine-needle aspiration on all such lesions; however, all positive lesions have to be resected, and negative cytologic findings in most cases have to be confirmed by histology on a resected specimen. While much has been made of sparing patients an operation, the high probability of finding a malignancy (particularly in smokers older than 35 years) and the excellent chance for surgical cure when the tumor is small both suggest an aggressive approach to these lesions.

The application of low-dose spiralCTscanning to high-risk populations is under investigation. The test identifies a large number of asymptomatic pulmonary nodules that require evaluation. Approximately 23% of screened high-risk patients have an abnormality, and ~12% of the detected abnormalities are lung cancer. Criteria for

distinguishing cancers from nonmalignant lesions short of a lobectomy are being developed. Lesions >1 cm are usually resected; those£1 cm are followed for change at 3-month intervals. Although a number of early lung cancers are detected in this way, it is not yet clear that survival is improved.

Radiotherapy with curative intent Patients with stage III disease, as well as patients with stage I or II disease who refuse surgery or are not candidates for pulmonary resection, should be considered for radiation therapy with curative intent. The decision to administer high-dose radiotherapy is based on the extent of disease and the volume of the chest that requires irradiation. Patients with distant metastases, malignant pleural effusion, or cardiac involvement are generally not considered for curative radiation treatment. The median survival period for patients with unresectable non-small cell lung cancer localized to the chest who undergo primary radiotherapy with curative intent is<1 year. However, 6% of these patients are alive at 5 years and are cured by radiotherapy alone. In addition to being potentially curative, radiotherapy, by controlling the primary tumor, may increase the quality and length of life of noncured patients. Treatment usually involves midplane doses of 55 to 60 Gy, and the major concern is the amount of lung parenchyma and other organs in the thorax included in the treatment plan, including the spinal cord, heart, and esophagus. In patients with a major degree of underlying pulmonary disease, the treatment plan may have to be compromised because of the deleterious effect of radiation on pulmonary function. The risk of radiation pneumonitis is proportional to the radiation dose and the volume of lung in the field. The full clinical syndrome (dyspnea, fever, and radiographic infiltrate corresponding to the treatment port) occurs in 5% of cases. Acute radiation esophagitis occurs during treatment but usually is self-limited, while spinal cord injury should be avoided by careful treatment planning. Continuous hyperfractionated accelerated radiation therapy (CHART) involves delivery of 36 treatments of 1.5 Gy given 3 times a day for 12 consecutive days to a total dose of 54 Gy. The 2-year survival rate increased from 20 to 29% with CHART, although more esophagitis occurred. Brachytherapy (local radiotherapy delivered by placing radioactive "seeds" in a catheter in the tumor bed) provides a way to give a high local dose while sparing surrounding normal tissue.

Combined-modality therapy with curative intent After apparently complete resection, adjuvant radiation therapy has not been shown to improve survival. Meta-analysis of studies with post-operative radiation therapy found it to be deleterious to survival in patients with stage I and II disease.

Carcinomas of the superior pulmonary sulcus producing *Pancoast's syndrome* are usually treated with combined radiotherapy and surgery. Patients with these carcinomas should have the usual preoperative staging procedures, including mediastinoscopy andCT scans to determine tumor extent and a neurologic examination (and sometimes nerve conduction studies) to document neurologic findings. Sometimes a histologic diagnosis is not made, but the combination of tumor location and pain distribution permit a diagnostic accuracy for cancer of>90%. If mediastinoscopy is negative, two curative approaches may be used in treating a Pancoast's syndrome tumor. Preoperative irradiation [30 Gy in 10 treatments] is given to the area, followed by an en bloc resection of the tumor and involved chest wall 3 to 6 weeks later. The 3 year survival rate is 42% for epidermoid and 21% for adeno- and large cell carcinomas. The second approach involves radiotherapy alone in curative doses and standard fractionation, which leads to

survival rates similar to those from combined-modality therapy.

A meta-analysis of chemotherapy in non-small cell lung cancer used updated data on 9387 individual patients from 52 randomized trials, both published and unpublished, with the main outcome measure being survival. Regimens containing cisplatin were significantly more effective than no treatment. Trials in early-stage disease comparing surgery with surgery plus chemotherapy gave a hazard ratio of 0.87 (13% reduction in risk of death at 5 years) in favor of chemotherapy. Confidence intervals of these data are wide. However, adjuvant chemotherapy is, in general, not considered standard treatment.

The most impressive benefits were obtained when chemotherapy was added to radiotherapy for locally advanced disease (stage IIIB and some stage IIIA disease) and when chemotherapy was given preoperatively in a neoadjuvant fashion in stage IIIA disease. Preoperative neoadjuvant chemotherapy is widely used for stage IIIA disease. Preoperative combined modality therapy followed by surgical resection has given promising early results. Whether the surgery adds benefit after chemoradiotherapy has not been defined. Provided the risk/benefit ratio of using chemotherapy is discussed appropriately with patients, such therapy can be given in a noninvestigational setting. For stage IIIA disease, resection followed by postoperative radiation plus chemotherapy for N2 disease, neoadjuvant chemotherapy followed by surgical resection, or neoadjuvant chemoradiotherapy followed by resection are options. For stage IIIB and bulky IIIA disease, neoadjuvant chemotherapy (2 or 3 cycles of a cisplatin-based combination) followed by chest radiation therapy (60 Gy) has improved median survival time from 10 to 14 months and the 5-year survival rate from 7 to 17% compared to results with radiation therapy alone. Administration of radiation and chemotherapy concurrently is being tested; myelotoxicity and esophagitis are increased, but survival improvement is not yet proven. Randomized clinical trials also are needed to evaluate the usefulness of the new agents with activity against non-small cell lung cancer, including the taxanes (paclitaxel and docetaxel), vinorelbine, gemcitabine, and camptothecins (topotecan and CPT-11) in both adjuvant and neoadjuvant settings.

Disseminated Non-Small Cell Lung Cancer The 70% of patients who have unresectable non-small cell cancer have a poor prognosis. Patients with performance status scores of 0 (asymptomatic), 1 (symptomatic, fully ambulatory), 2 (in bed <50% of the time), 3 (in bed >50% of the time), and 4 (bedridden) have median survival times of 34, 25, 17, 8, and 4 weeks, respectively. Standard medical management, the judicious use of pain medications, the appropriate use of radiotherapy, and outpatient chemotherapy form the cornerstone of management. Patients whose primary tumor is causing symptoms such as bronchial obstruction with pneumonitis, hemoptysis, or upper airway or superior vena cava obstruction should have radiotherapy to the primary tumor. The case for prophylactic treatment of the asymptomatic patient is to prevent major symptoms from occurring in the thorax. However, if the patient can be followed closely, it may be appropriate to defer treatment until symptoms develop. Usually a course of 30 to 40 Gy over 2 to 4 weeks is given to the tumor. Radiation therapy provides relief of intrathoracic symptoms with the following frequencies: hemoptysis, 84%; superior vena cava syndrome, 80%; dyspnea, 60%; cough, 60%; atelectasis, 23%; and vocal cord paralysis, 6%. Cardiac tamponade (treated with pericardiocentesis and radiation therapy to the heart), painful bony metastases (with relief in 66%), brain or

spinal cord compression, and brachial plexus involvement may also be palliated with radiotherapy. Usually, with brain metastases and cord compression, dexamethasone (25 to 100 mg/d in four divided doses) is also given and then rapidly tapered to the lowest dosage that relieves symptoms.

Brain metastases often are isolated instances of relapse in patients with adenocarcinoma of the lung otherwise controlled by surgery or radiotherapy. However, there is no proven value for prophylactic cranial irradiation or forCT scans of the head in asymptomatic patients.

Pleural effusions are common and are usually treated with thoracentesis. If they recur and are symptomatic, chest tube drainage with a sclerosing agent such as intrapleural talc is used. First, the chest cavity is completely drained. Xylocaine 1% is instilled (15 mL), followed by 50 mL normal saline. Then, 10 g sterile talc is dissolved in 100 mL normal saline, and this solution is injected through the chest tube. The chest tube is clamped for 4 h if tolerated, and the patient is rotated onto different sides to distribute the sclerosing agent. The chest tube is removed 24 to 48 h later, after drainage has become slight (usually <100 mL/24 h).VATS has been used to drain and treat large malignant effusions. Symptomatic endobronchial lesions that recur after surgery or radiotherapy or that develop in patients with severely compromised pulmonary function are difficult to treat with conventional therapy. Neodynium-YAG (yttrium-aluminum-garnet) laser therapy administered through a flexible fiberoptic bronchoscope (usually under general anesthesia) can provide palliation in 80 to 90% of patients even when the tumor has relapsed after radiotherapy. Local radiotherapy delivered by brachytherapy, photodynamic therapy using a photosensitizing agent, and endobronchial stents are other measures that can relieve airway obstruction from tumor.

Chemotherapy The use of chemotherapy for non-small cell lung cancer requires careful judgment to balance potential benefits and toxicity. Modest survival benefits (of 1 to 2 months), symptom palliation, and improved quality of life may accrue from combination chemotherapy. Randomized trials in advanced disease comparing supportive care with supportive care plus chemotherapy gave a hazard ratio of 0.73 (27% reduction in risk of death at 1 year) in favor of including chemotherapy. Economic analysis has found chemotherapy to be cost-effective palliation. Combination chemotherapy produces an objective tumor response in ~30 to 40% of patients; the response is complete in <5%. Median survival for chemotherapy-treated patients is 9 to 10 months, and the 1-year survival rate is 40%. Thus, in patients with non-small cell lung cancer who desire chemotherapy, it is reasonable to give chemotherapy if the patient is ambulatory, has not received prior chemotherapy, and is able to understand and accept the risk/benefit ratio from such therapy. The chemotherapy should be one of the published standard regimens, such as paclitaxel plus carboplatin, paclitaxel plus cisplatin, or vinorelbine plus cisplatin. Improved antiemetics have made treatment tolerable on an outpatient basis. New drugs with proven activity in non-small cell lung cancer include docetaxel, irinotecan, and gemcitabine. All eligible patients should be encouraged to enter clinical studies that are designed to determine the benefits and toxicities of these new treatments.

Small Cell Lung Cancer Untreated patients with small cell lung cancer have a median survival period of 6 to 17 weeks, while patients treated with combination chemotherapy

have a median survival period of 40 to 70 weeks. Thus, chemotherapy with or without radiotherapy or surgery can prolong survival in patients with small cell lung cancer. The goal of treatment is to achieve a complete clinical regression of tumor documented by repeating the initial positive staging procedures, particularly fiberoptic bronchoscopy with washings and biopsy. The initial response, determined 6 to 12 weeks after the start of therapy, predicts both the median and long-term survivals and the potential for cure. Patients who achieve a complete clinical regression survive longer than patients with only partial regression, who in turn survive longer than patients with no response. Complete response is required for long-term (>3-year) survival.

After initial staging, patients are classified as having limited or extensive disease and as being physiologically able or not able to tolerate combination chemotherapy or chemoradiotherapy. The overall mortality rate from initial combination chemotherapy even in these selected patients is 1 to 5%, comparable with the operative mortality rate for pulmonary resection. Such therapy should be reserved for ambulatory patients with no prior chemotherapy or radiotherapy; no other major medical problems; and adequate heart, liver, renal, and bone marrow function. The arterial P_{O2} on room air should be >6.6 kPa (50 mmHg), and there should be no CO_2retention. For patients with limitations in any of these areas, the initial combined-modality therapy or chemotherapy must be modified to prevent undue toxicity. In all patients, these treatments must be coupled with supportive care for infectious, hemorrhagic, and other medical complications.

Chemotherapy The combination most widely used is etoposide plus cisplatin or carboplatin, given every 3 weeks on an outpatient basis for 4 to 6 cycles. Another active regimen is etopside, cisplatin, and paclitaxel. Increased dose intensity of chemotherapy adds toxicity without clear survival benefit. Appropriate supportive care (antiemetic therapy, administration of fluid and saline boluses with cisplatin, monitoring of blood counts and blood chemistries, monitoring for signs of bleeding or infection, and, as required, administration of erythropoietin and granulocyte colony-stimulating factor) and adjustment of chemotherapy doses on the basis of nadir granulocyte counts are essential. The initial combination chemotherapy may result in moderate to severe granulocytopenia (e.g., granulocyte counts <500 to 1500/uL) and thrombocytopenia (platelet counts <50,000 to 100,000/uL). After the initial 4 to 6 cycles of therapy, patients should be restaged to determine if they have entered a complete clinical remission, indicated by complete disappearance of all clinically evident lesions and paraneoplastic syndromes, or a partial remission, or have no response or tumor progression (seen in 10 to 20% of patients). Chemotherapy is then stopped in responding patients. More prolonged chemotherapy has not been shown to be of value. Patients whose tumors are progressing or not responding should be switched to a new, experimental chemotherapy regimen. Oral etoposide, as a single agent, has been shown to be of clinical benefit in the initial treatment of patients who are elderly or have a very poor performance status.

Radiotherapy High-dose (40 Gy) radiotherapy to the whole brain should be given to patients with documented brain metastases. Prophylactic cranial irradiation (PCI) may be given to patients with complete responses, since it significantly decreases the development of brain metastases (which occur in 60 to 80% of patients living³2 years who do not receive PCI), but survival benefit is small (5%). Because some studies indicate possible deficits in cognitive ability that could be related to PCI, the long-term quality of life after PCI needs to be further studied. The patient needs to be informed of

the risks and benefits. In the case of symptomatic, progressive lesions in the chest or at other critical sites, if radiotherapy has not yet been given to these areas, it may be administered in full doses (e.g., 40 Gy to the chest tumor mass).

Combined-modality therapy Most patients with limited-stage small cell lung cancer should receive combined-modality therapy with etoposide plus cisplatin (or other platinum-containing regimen) and concurrent chest radiotherapy. Acute and chronic toxicities are expected with chemoradiotherapy, particularly when the chemotherapy and radiotherapy are given concurrently. However, the addition of chest radiation therapy to chemotherapy reduces the local failure rate and improves survival. Patients should be selected (limited-stage disease, a performance status of 0 to 1, and initial good pulmonary function) such that radiotherapy can be given in full doses and in a manner that does not sacrifice too much lung function. Some studies show twice-daily radiation fractions produce less toxicity and improve survival compared to once-daily treatments, but large randomized trials are still needed.

For extensive-stage disease, initial chest radiotherapy usually is not advocated. However, for favorable patients (e.g., those with a performance status of 0 to 1, good pulmonary function, and only one site of extensive disease), the addition of chest radiotherapy to chemotherapy can be considered. For all patients, if chemotherapy is inadequate to relieve local tumor symptoms, a course of radiotherapy can be added.

About 20 to 30% of patients with limited-stage disease and 1 to 5% of patients with extensive-stage disease are cured. About 50% of patients with limited-stage and 30% of patients with extensive-stage disease enter complete remission, and 90 to 95% of all patients have complete or partial responses. These responses increase the median survival period to 10 to 12 months for patients with extensive-stage disease and to 14 to 18 months for patients with limited-stage disease, as compared with 2 to 4 months for untreated patients. In addition, most patients have relief of their tumor-related symptoms and improvement of performance status. However, the maintenance of good performance status in a patient receiving outpatient chemotherapy requires judgment and skill to avoid undue therapeutic toxicity. New treatments, such as new drug combinations, very intensive initial or "reinduction" therapy with autologous bone marrow infusion, and novel ways of combining chemotherapy, radiotherapy, and surgery, should be given only in the context of an approved clinical protocol.

Although surgical resection is not routinely recommended for small cell lung cancer, occasional patients meet the usual requirements for resectability (stage I or II disease with negative mediastinal nodes). Moreover, this histologic diagnosis is made in some patients only on review of the resected surgical specimen. Such patients have been reported to have high cure rates (>25%) if adjuvant chemotherapy is used.

LUNG CANCER PREVENTION

Deterring children from taking up smoking is likely to be the most effective lung cancer prevention. Smoking cessation programs are successful in 5 to 20% of volunteers; the poor efficacy is because of the nature of nicotine addiction. Early diagnosis strategies have the problem of high false-positive rates, which add to the expense and the failure of such strategies to result in improved survival.

Chemoprevention may be an approach to reduce lung cancer risk. Patients with head and neck cancer, who are at increased risk of developing lung cancer, experienced a decrease in second cancers when given 13-cis retinoic acid. However, the drug causes significant toxicity, and its activity is not yet confirmed. Vitamin E andb-carotene supplements actually increase the risk of lung cancer. Thus, currently no strategy for chemoprevention of lung cancer has been proven effective.

BENIGN LUNG NEOPLASMS

The benign neoplasms of the lung, representing <5% of all primary tumors, include bronchial adenomas and hamartomas (90% of such lesions) and a group of very uncommon neoplasms (chondromas, fibromas, lipomas, hemangiomas, leiomyomas, teratomas, pseudolymphomas, and endometriosis). The diagnostic and primary-treatment approach is basically the same for all these neoplasms. They can present as central masses causing airway obstruction, cough, hemoptysis, and pneumonitis. The masses may or may not be visible on radiographs but usually are accessible to fiberoptic bronchoscopy. Alternatively, they can present without symptoms as solitary pulmonary nodules and thus will be evaluated as part of a solitary pulmonary nodule workup. In all cases, the extent of surgery must be determined at operation, and a conservative procedure with appropriate reconstructions is usually performed.

BRONCHIAL ADENOMAS

Bronchial adenomas (80% are central) are slow-growing, endobronchial lesions; they represent 50% of all benign pulmonary neoplasms. About 80 to 90% are carcinoids, 10 to 15% are adenocystic tumors (or cylindromas), and 2 to 3% are mucoepidermoid tumors. Adenomas present in patients 15 to 60 years old (average age, 45) as endobronchial lesions and are often symptomatic for several years. Patients may have a chronic cough, recurrent hemoptysis, or obstruction with atelectasis, lobar collapse, or pneumonitis and abscess formation. Bronchial carcinoids, which usually follow a benign course, and small cell lung cancers, which are highly malignant, both express a neuroendocrine phenotype similar to the Kulchitsky cell. This cell is part of the amine precursor uptake and decarboxylation (APUD) system. Carcinoids, like small cell lung cancers, may secrete other hormones, such asACTH or arginine vasopressin, and can cause paraneoplastic syndromes that resolve on resection. In addition, bronchial carcinoid metastases (usually to the liver) may produce the carcinoid syndrome, with cutaneous flush, bronchoconstriction, diarrhea, and cardiac valvular lesions (Chap. 93), which small cell lung cancer does not. Occasionally, pathologists may have difficulty distinguishing carcinoids from small cell lung cancers. Carcinoid tumors that have an unusually aggressive histologic appearance (referred to as *atypical carcinoids*) metastasize in 70% of cases to regional nodes, liver, or bone, compared with only a 5% rate of metastasis for carcinoids with typical histology.

Bronchial adenomas of all types, because of their endobronchial and often central location, are usually visible by fiberoptic bronchoscopy; and tissue for histologic diagnosis is obtained in this manner. Because they are hypervascular, they can bleed profusely after bronchoscopic biopsy, and this problem should be anticipated. Bronchial adenomas must be dealt with as potentially malignant and thus require removal not only

for symptom relief but also because they can be locally invasive or recurrent, potentially can metastasize, and may produce paraneoplastic syndromes. Surgical excision is the primary treatment for all types of bronchial adenomas. The extent of surgery is determined at operation and should be as conservative as possible. Often bronchotomy with local excision, sleeve resection, segmental resection, or lobectomy is sufficient. Five-year survival rates after surgical resection are 95%, decreasing to 70% if regional nodes are involved. The treatment of metastatic pulmonary carcinoids is unclear because they can either be indolent or behave more like small cell lung carcinoma. Assessment of the tempo and histology of the disease in the individual patient is necessary to determine if and when chemotherapy or radiotherapy is indicated.

HAMARTOMAS

Pulmonary hamartomas have a peak incidence at age 60 and are more frequent in men than in women. Histologically, they contain normal pulmonary tissue components (smooth muscle and collagen) in a disorganized fashion. They are usually peripheral, clinically silent, and benign in their behavior. Unless the radiographic findings are pathognomonic for hamartoma, with "popcorn" calcification, the lesions usually have to be resected for diagnosis, particularly if the patient is a smoker.VATS may minimize the surgical complications.

METASTATIC PULMONARY TUMORS

The lung is a frequent site of metastases from primary cancers outside the lung. Usually such metastatic disease is considered incurable. However, two special situations should be borne in mind. The first is the development of a solitary pulmonary shadow on a chest x-ray in a patient known to have an extrathoracic neoplasm. This shadow may represent a metastasis or a new primary lung cancer. Because the natural history of lung cancer is often worse than that of other primary tumors, a single pulmonary nodule in a patient with a known extrathoracic tumor is approached as though the nodule is a primary lung cancer, particularly if the patient is older than 35 years and a smoker. If a vigorous search for other sites of active cancer proves negative, the nodule is surgically resected. Second, in some cases, multiple pulmonary nodules can be resected with curative intent. This tactic is usually recommended if, after careful staging, it is found that (1) the patient can tolerate the contemplated pulmonary resection, (2) the primary tumor has been definitively and successfully treated, and (3) all known metastatic disease can be encompassed by the projected pulmonary resection. The key is selection and screening of patients to exclude those with uncontrolled primary tumors and extrapulmonary metastases. Primary tumors whose pulmonary metastases have been successfully resected for cure include osteogenic and soft tissue sarcomas; colon, rectal, uterine, cervix, and corpus tumors; head and neck, breast, testis, and salivary gland cancer; melanoma; and bladder and kidney tumors. Five-year survival rates of 20 to 30% have been found in carefully selected patients, and dramatic results have been achieved in patients with osteogenic sarcomas, where resection of pulmonary metastases (sometimes requiring several thoracotomies) is becoming a standard curative treatment approach.

(Bibliography omitted in Palm version)

Back to Table of Contents

89. BREAST CANCER - *Marc E. Lippman*

Breast cancer is a malignant proliferation of epithelial cells lining the ducts or lobules of the breast. In the year 2000, about 185,000 cases of invasive breast cancer and 42,000 deaths occurred in the United States. Mortality from breast cancer has begun to decrease. Epithelial malignancies of the breast are the most common cause of cancer in women (excluding skin cancer), accounting for about one-third of all cancer in women. This chapter will not consider rare malignancies of the breast, including sarcomas and lymphomas. Human breast cancer is a clonal disease; a single transformed cell -- the end result of a series of somatic (acquired) or germline mutations -- is able to express full malignant potential. Thus, breast cancer may exist for a long period as either a noninvasive disease or an invasive but nonmetastatic disease.

GENETIC CONSIDERATIONS

Not more than 10% of human breast cancers can be linked directly to germline mutations. Several genes have been implicated in familial cases. The Li-Fraumeni syndrome is characterized by inherited mutations in the p53 tumor suppressor gene, which lead to an increased incidence of breast cancer, osteogenic sarcomas, and other malignancies.

Another putative tumor suppressor gene, *BRCA-1*, has been identified at the chromosomal locus 17q21; this gene encodes a zinc finger protein, and the product therefore may function as a transcriptional factor. The gene appears to be involved in gene repair. Women who inherit a mutated allele of this gene from either parent have an approximately 60 to 80% lifetime chance of developing breast cancer and about a 33% chance of developing ovarian cancer. Men who carry a mutant allele of the gene have an increased incidence of prostate cancer but usually not of breast cancer. A third gene, termed *BRCA-2*, which has been localized to chromosome 13q12, is associated with an increased incidence of breast cancer in men and women.

BRCA-1 and *BRCA-2* can now be sequenced readily and germline mutations detected; patients with these mutations can be counseled appropriately. All women with strong family histories for breast cancer should be referred to genetic screening programs whenever possible, particularly women of Ashkenazi Jewish descent who have a high likelihood of a specific *BRCA-1* mutation (deletion of adenine and guanine at position 185).

Even more important than the role these genes play in inherited forms of breast cancer may be their role in sporadic breast cancer. The p53 mutation is present in approximately 40% of human breast cancers as an acquired defect. Evidence for *BRCA-1* mutation in primary breast cancer has not been reported. However, decreased expression of *BRCA-1* mRNA and abnormal cellular location of the *BRCA-1* protein have been found in some breast cancers. Loss of heterozygosity of some genes suggests that tumor-suppressor activity may be inactivated in sporadic cases of human breast cancer. Finally, one dominant oncogene plays a role in about a quarter of human breast cancer cases. The product of this gene, a member of the epidermal growth factor receptor superfamily, is called *erbB2* (HER-2, neu) and is overexpressed in these breast cancers due to gene amplification; this overexpression can transform human breast

epithelium.

EPIDEMIOLOGY

Breast cancer is a hormone-dependent disease. Women without functioning ovaries who never receive estrogen replacement do not develop breast cancer. The female to male ratio is about 150 to 1. For most epithelial malignancies, a log-log plot of incidence versus age shows a straight-line increase with every year of life. A similar plot for breast cancer shows the same straight-line increase but with a decrease in slope beginning at the age of menopause. The three dates in a woman's life that have a major impact on breast cancer incidence are age at menarche, age at first full-term pregnancy, and age at menopause. Women who experience menarche at age 16 have only 50 to 60% of the breast cancer risk of a woman having menarche at age 12; the lower risk persists throughout life. Similarly, menopause occurring 10 years before the median age of menopause (52 years), whether natural or surgically induced, reduces lifetime breast cancer risk by about 35%. Women who have a first full-term pregnancy by age 18 have a reduced (30 to 40%) risk of breast cancer compared with nulliparous women. Thus, length of menstrual life -- particularly the fraction occurring before first full-term pregnancy -- is a substantial component of the total risk of breast cancer. This factor can account for 70 to 80% of the variation in breast cancer frequency in different countries.

International variation in incidence has provided some of the most important clues on hormonal carcinogenesis. A woman living to age 80 in North America has one chance in nine of developing invasive breast cancer. Asian women have one-fifth to one-tenth the risk of breast cancer of women in North America or Western Europe. Asian women have substantially lower concentrations of estrogens and progesterone. These differences cannot be explained on a genetic basis, because Asian women living in a western environment have sex steroid hormone concentrations and risk identical to that of their western counterparts. These women also differ markedly in height and weight from Asian women in Asia; height and weight are critical regulators of age of menarche and have substantial effects on plasma concentrations of estrogens.

The role of diet in breast cancer etiology is controversial. While there are associative links between total caloric and fat intake and breast cancer risk, the exact role of fat in the diet is unproven. However, there is a risk associated with moderate alcohol intake; the mechanism is unknown. Recommendations favoring abstinence from alcohol must be weighed against other social pressures and the possible cardioprotective effect of moderate alcohol intake.

The potential role of exogenous hormones in breast cancer is of extraordinary importance, because millions of American women regularly use oral contraceptives and postmenopausal hormone replacement therapy (HRT). The most credible meta-analyses of oral contraceptive use suggest that these agents cause little if any increased risk of breast cancer. By contrast, oral contraceptives offer a substantial protective effect against ovarian epithelial tumors and endometrial cancers. Far more controversial are the data surrounding HRT in hypogonadal and/or menopausal women. First, HRT with estrogens alone, usually in the form of equine conjugated estrogens, provides less than the physiologic equivalent of premenopausal estrogens but is

associated with an increased risk of endometrial cancer, a reduction in the symptoms of estrogen deprivation, a reduction in osteoporosis and resultant hip fractures, and a one-third reduction in deaths due to cardiovascular disease. Meta-analyses suggest a small increase in breast cancer incidence, particularly with high dosages and a long duration of treatment. For the average woman, the negative effect on the breast is probably outweighed by protective effects on bone and heart. Preliminary data suggest that there is a reduction in the risk of colon cancer as well.

The addition of progestogens toHRTregimens drastically reduces the risk of endometrial cancer. It is not clear whether the protective effects against cardiovascular and osteoporotic bone diseases are altered. However, progestogens are copromoters of breast cancer in model systems, and an increased risk of breast cancer is possible.

Whether a history of previous biopsy findings of atypical hyperplasia or in situ carcinoma or strong family histories of breast cancer alter the risk-to-benefit ratios forHRT is unknown. It is likely that the average woman benefits from HRT. The risks of HRT in patients with a positive family history and patients with a remote personal history of breast cancer are unknown.

In addition to the other factors, radiation may be a risk factor in younger women. Women who have been exposed before age 30 to radiation in the form of multiple fluoroscopies (200 to 300 cGy) or treatment for Hodgkin's disease (>3600 cGy) have a substantial increase in risk of breast cancer, whereas radiation exposure after age 30 appears to have a minimal carcinogenic effect on the breast.

EVALUATION OF BREAST MASSES IN MEN AND WOMEN

Because the breasts are a common site of potentially fatal malignancy in women and because they frequently provide clues to underlying systemic diseases in both men and women, examination of the breast is an essential part of the physical examination. Unfortunately, internists frequently do not examine the breast in men, and, in women, they are apt to refer this evaluation to gynecologists. Because of the association between early detection and improved outcome, it is the duty of every physician to distinguish breast abnormalities at the earliest possible stage and to institute a definite diagnostic workup. It is for this reason that all women should be trained in self-examination of the breasts. Although breast cancer in men is unusual, unilateral lesions should be evaluated in the same manner as in women, with the recognition that gynecomastia in men can sometimes begin unilaterally and is often asymmetric. Nevertheless, about as many suspicious breast lesions are now detected by screening mammography as by physical examination.

Virtually all breast cancer is diagnosed by biopsy of a nodule detected either on a mammogram or by palpation. Algorithms have been developed to enhance the likelihood of diagnosing breast cancer and reduce the frequency of unnecessary biopsy.

The Palpable Breast Mass Women should be strongly encouraged to examine their breasts monthly. The minimum benefit of this practice is the greater likelihood of detecting a mass at a smaller size, when it can be treated with more limited surgery. Breast examination by the physician should be performed in good light so as to see

retractions and other skin changes. The nipple and areolae should be inspected, and an attempt should be made to elicit nipple discharge. All regional lymph node groups should be examined, and any lesions should be measured. While lesions with certain features are more likely to be cancerous (hard, irregular, tethered or fixed, or painless lesions), physical examination alone cannot exclude malignancy. Furthermore, a negative mammogram in the presence of a persistent lump in the breast does not exclude malignancy.

In premenopausal women, lesions that are either equivocal or nonsuspicious on physical examination should be reexamined in 2 to 4 weeks, during the follicular phase of the menstrual cycle. Days 5 to 7 of the cycle are the best time for breast examination. A dominant mass in a postmenopausal woman or a dominant mass that persists through a menstrual cycle in a premenopausal woman should be aspirated by fine-needle biopsy or referred to a surgeon. If nonbloody fluid is aspirated and the lesion is thereby cured, the diagnosis (cyst) and therapy have been accomplished together. Solid lesions that are persistent, recurrent, complex or bloody cysts require mammography and biopsy, although in selected patients the so-called triple diagnostic techniques (palpation, mammography, aspiration) can be used to avoid biopsy (Figs. 89-1,89-2, and89-3). Ultrasound can be used in place of fine-needle aspiration to distinguish cysts from solid lesions. Not all solid masses are detected by ultrasound; thus, a palpable mass that is not visualized on ultrasound must be presumed to be solid.

Several points are essential in pursuing these management decision trees. First, risk factor analysis is not part of the decision structure. Second, fine-needle aspiration should be used only in centers that have proven skill in obtaining such specimens and analyzing them. Although the likelihood of cancer is low in the setting of a "triple negative" (benign-feeling lump, negative mammogram, and negative fine-needle aspiration), it is not zero, and the patient and physician must be aware of about a 1% risk of false negativity. Third, additional technologies such as magnetic resonance imaging, ultrasound, and sestamibi imaging cannot be used to exclude the need for biopsy, although in unusual circumstances they may provoke a biopsy.

The Abnormal Mammogram Screening mammography has reduced the lethality of breast cancer by promoting detection at an earlier stage. The procedure is justified on an annual basis for women over age 40.

Screening mammography should not be confused with diagnostic mammography, which is performed after a palpable abnormality has been detected. Diagnostic mammography is aimed at evaluating the rest of the breast before biopsy is performed, or occasionally is part of the triple test strategy to exclude immediate biopsy.

Subtle abnormalities that are first detected by screening mammography should be evaluated carefully by compression or magnified views. These abnormalities include clustered microcalcifications, densities (especially if spiculated), and new or enlarging architectural distortion. For some nonpalpable lesions ultrasound may be helpful either to identify cysts or to guide biopsy. If there is no palpable lesion and detailed mammographic studies are unequivocally benign, the patient should have routine follow-up appropriate to the patient's age.

If a nonpalpable mammographic lesion has a low index of suspicion, mammographic follow-up in 3 to 6 months is reasonable. Workup of indeterminate and suspicious lesions has been rendered more complex by the advent of stereotactic biopsies. Morrow and colleagues have suggested that these procedures are indicated for lesions that require biopsy but are likely to be benign -- that is, for cases in which the procedure probably will eliminate additional surgery. When a lesion is more probably malignant, open excisional biopsy should be performed with a needle localization technique. Others have proposed more widespread use of stereotactic core biopsies for nonpalpable lesions, on economic grounds and because diagnosis leads to earlier treatment planning. However, stereotactic diagnosis of a malignant lesion does not eliminate the need for definitive surgical procedures, particularly if breast conservation is attempted. For example after a breast biopsy with needle localization (i.e., local excision) of a stereotactically diagnosed malignancy, reexcision may still be necessary to achieve negative margins. To some extent, these issues are decided on the basis of referral pattern and the availability of the resources for stereotactic core biopsies. A reasonable approach is shown in Fig. 89-4.

Breast Masses in the Pregnant or Lactating Woman During pregnancy, the breast grows under the influence of estrogen, progesterone, prolactin, and human placental lactogen. Lactation is suppressed by progesterone, which blocks the effects of prolactin. After delivery, lactation is promoted by the fall in progesterone levels, which leaves the effects of prolactin unopposed. The development of a dominant mass during pregnancy or lactation should never be attributed to hormonal changes, and biopsy should never be performed under local anesthesia. Breast cancer develops in 1 in every 3000 to 4000 pregnancies. Stage for stage, breast cancer in pregnant patients is no different from premenopausal breast cancer in nonpregnant patients. However, pregnant women often have more advanced disease because a breast mass was ignored.

Benign Breast Masses Only about 1 in every 5 to 10 breast biopsies leads to a diagnosis of cancer, although the rate of positive biopsies varies in different countries. (These differences may be related to interpretation and availability of mammograms.) The vast majority of benign breast masses are due to "fibrocystic" disease, a descriptive term for small fluid-filled cysts and modest epithelial cell and fibrous tissues hyperplasia. However, fibrocystic disease is a histologic, not a clinical, diagnosis, and women who have had a biopsy with benign findings are at greater risk of developing breast cancer than those who have not had a biopsy. The subset of women with ductal or lobular cell proliferation (about 30% of patients), particularly the small fraction (3%) with atypical hyperplasia, have a fourfold greater risk of developing breast cancer than unbiopsied women, and the increase in the risk is about ninefold for women in this category who also have an affected first-degree relative. Thus, careful follow-up of these patients is required. By contrast, patients with a benign biopsy without atypical hyperplasia are at little risk and may be followed routinely.

SCREENING

Breast cancer is virtually unique among the epithelial tumors in adults in that screening (in the form of annual mammography) has been proven to improve survival. Meta-analysis examining outcomes from every randomized trial of mammography conclusively shows a 25 to 30% reduction in the chance of dying from breast cancer

with annual screening after age 50; the data for women between ages 40 and 50 are almost as positive. It seems prudent to recommend annual mammography for women past the age of 40. Although no randomized study of breast self-examination (BSE) has ever shown any improvement in survival, its major benefit appears to be identification of tumors appropriate for conservative local therapy. Better mammographic technology, including digitized mammography, routine use of magnified views, and greater skill in mammographic interpretation, combined with newer diagnostic techniques (magnetic resonance imaging, magnetic resonance spectroscopy, positron emission tomography, etc.) may make it possible to identify breast cancers yet more reliably and earlier.

STAGING

Correct staging of breast cancer patients is of extraordinary importance. Not only does it permit an accurate prognosis, but in many cases therapeutic decision-making is based largely on the TNM classification (Table 89-1). Comparison with historic series should be undertaken with caution, as the staging has changed in the past 10 years.

TREATMENT

Primary Breast Cancer A series of randomized clinical trials both in the United States and abroad have shown that breast-conserving treatments, consisting of the removal of the primary tumor by some form of lumpectomy with or without irradiating the breast, results in a survival that is as good as that after extensive procedures, such as mastectomy or modified radical mastectomy, with or without further irradiation. While breast conservation is associated with a possibility of recurrence in the breast, 10-year survival is at least as good as that after more radical surgery. Postoperative radiation to regional nodes following mastectomy is also associated with an improvement in survival. Since radiation therapy can also reduce the rate of local or regional recurrence, it should be strongly considered following mastectomy for women with high-risk primary tumors (i.e., T2 in size, positive margins, positive nodes). At present, approximately one-third of women in the United States are managed by lumpectomy.
Breast-conserving surgery is not suitable for all patients; it is not generally suitable for tumors >5 cm (or for smaller tumors if the breast is small), for tumors involving the nipple areola complex, for tumors with extensive intraductal disease involving multiple quadrants of the breast, for women with a history of collagen-vascular disease, and for women who either do not have the motivation for breast conservation or do not have convenient access to radiation therapy. However, these groups probably do not account for more than one-third of patients. Thus, a great many women who undergo mastectomy could safely avoid this procedure.

An extensive intraductal component is a predictor of recurrence in the breast, and so are several clinical variables. Both axillary lymph node involvement and involvement of vascular or lymphatic channels by metastatic tumor in the breast are associated with a higher risk of relapse in the breast but are not contraindications to breast-conserving treatment. When these patients are excluded, and when lumpectomy with negative tumor margins is achieved, breast conservation is associated with a recurrence rate in the breast of less than 10%. The survival of patients who have recurrence in the breast is somewhat worse than that of women who do not. Thus, recurrence in the breast is a negative prognostic variable for long-term survival. However, recurrence in the breast is

not the *cause* of distant metastasis. If recurrence in the breast caused metastatic disease, then women treated with lumpectomy, who have a higher rate of recurrence in the breast, should have poorer survival. Most patients should consult with a radiation oncologist before making a final decision concerning local therapy. However, a multimodality clinic approach in which the surgeon, radiation oncologist, medical oncologist, and other caregivers cooperate to evaluate the patient and develop a treatment is usually considered a major advantage by patients.

Adjuvant Therapy One of the significant advances in the treatment of solid tumors of adults has been the improved survival resulting from the use of systemic therapy after local management of breast cancer. More than one-third of the women who would otherwise die of metastatic breast cancer remain disease-free when treated with the appropriate systemic regimen.

PROGNOSTIC VARIABLES The most important prognostic variables are provided by *tumor staging*. The size of the tumor and the status of the axillary lymph nodes provide reasonably accurate information on the likelihood of tumor relapse. The relation of pathologic stage to 5-year survival is shown inTable 89-2. For most women, the need for adjuvant therapy can be readily defined on this basis alone. In the absence of lymph node involvement, involvement of microvessels (either capillaries or lymphatic channels) in tumors is nearly equivalent to lymph node involvement. The greatest controversy concerns women with intermediate prognoses. *There is no justification for adjuvant chemotherapy in women with tumors<1 cm in size whose axillary lymph nodes are negative*.

Other prognostic variables have been sought and some appear to influence disease-free and overall survival. What is less clear is whether they add to the information from pathologic staging.

Estrogen and progesterone receptor status are of prognostic significance. Tumors that lack either or both of these receptors are more likely to recur than tumors that have them.

Several *measures of tumor growth rate* correlate with early relapse. S-phase analysis using flow cytometry is the most accurate measure, and the indirect S-phase assessments using antigens associated with the cell cycle, such as PCNA (Ki67), are also valuable. Several studies suggest that tumors with a high proportion (more than the median) of cells in the S phase pose a greater risk of relapse and that chemotherapy offers the greatest survival benefit for these tumors. For this reason, some clinicians use S-phase assessment as a deciding factor for instituting adjuvant therapy when other pathologic features are unclear. Assessment of DNA content in the form of ploidy is of modest value, with nondiploid tumors having a somewhat worse prognosis.

Histologic classification of the tumor has also been used as a prognostic factor. Tumors with a poor nuclear grade have a higher risk of recurrence than tumors with a good nuclear grade. Semiquantitative measures such as the Elston score improve the reproducibility of this measurement.

Molecular changes in the tumor are also useful. Tumors that overexpress erbB2

(HER-2/neu) or have a mutated p53 gene have a worse prognosis. Particular interest has centered on erbB2 overexpression as measured by histochemistry. Tumors that overexpress erbB2 are more likely to respond to higher doses of doxorubicin-containing regimens. For this reason, erbB2 expression is usually worth measuring as a means of deciding on therapy.

To grow, a tumor must generate a neovasculature (Chap. 83). The presence of more microvessels in a tumor is associated with a worse prognosis.

Other variables that have also been used to evaluate prognosis include proteins associated with invasiveness, such as type IV collagenase, cathepsin D, plasminogen activator, plasminogen activator receptor, and the metastasis suppressor gene, nm23. None of these has been widely accepted as a prognostic variable for therapeutic decision-making. One problem in interpreting these prognostic variables is that most of them have not been examined in a study using a large cohort of patients.

ADJUVANT REGIMENS Selection of appropriate adjuvant chemotherapy or hormone therapy regimens is a highly controversial issue in some situations. Meta-analyses have helped to define broad limits for therapy but do not help in choosing optimal regimens or in choosing a regimen for certain subgroups of patients. A summary of recommendations is shown in Table 89-3. In general, premenopausal women for whom any form of adjuvant systemic therapy is indicated should receive chemotherapy for 6 months. The antiestrogen (tamoxifen) improves survival in premenopausal patients with positive estrogen receptor values and should be added following completion of chemotherapy. Prophylactic castration may also be associated with a substantial survival benefit (primarily in estrogen receptor-positive patients) but is not widely used in this country.

Data on postmenopausal women are also controversial. The impact of adjuvant chemotherapy is less clear-cut than in premenopausal patients, although some survival advantage has been shown. The first decision is whether chemotherapy or tamoxifen should be used. While adjuvant tamoxifen improves survival regardless of axillary lymph node status, the improvement in survival is modest for patients in whom multiple lymph nodes are involved. For this reason, it has been usual to give chemotherapy to postmenopausal patients who have no medical contraindications and who have more than one positive lymph node; tamoxifen is commonly given simultaneously or subsequently. For postmenopausal women for whom systemic therapy is warranted but who have a more favorable prognosis, tamoxifen may be used as a single agent.

Most comparisons of adjuvant chemotherapy regimens show little difference among them, although slight advantages for doxorubicin-containing regimens are usually seen.

One approach -- so-called neoadjuvant chemotherapy -- involves the administration of adjuvant therapy before definitive surgery and radiation therapy. Because the objective response rates of patients with breast cancer to systemic therapy in this setting exceed 75%, many patients will be "downstaged" and may become candidates for breast-conserving therapy. At least one large randomized study has failed to show any difference in survival using this approach.

Other adjuvant treatments under investigation include the use of new drugs, such as paclitaxel, and therapy based on alternative kinetic and biologic models. In such approaches, high doses of single agents are used separately in relatively dose-intensive cycling regimens. One large randomized trial for node-positive patients suggests that patients treated with doxorubicin-cyclophosphamide for four cycles followed by four cycles of paclitaxel have a substantial additional gain in survival as compared with women receiving doxorubicin-cyclophosphamide alone. Very high dose therapy with stem cell transplantation in the adjuvant setting has not proved superior.

Systemic Therapy of Metastatic Disease Nearly half of patients treated for apparently localized breast cancer develop metastatic disease. Although some of these patients can be salvaged by combinations of systemic and local therapy, most eventually succumb. Soft tissue, bony, and visceral (lung and liver) metastases each account for approximately one-third of sites of initial relapses. However, by the time of death, most patients will have bony involvement. Recurrences can appear at any time after primary therapy. Half of all initial cancer recurrences occur more than 5 years following initial therapy.

Because this diagnosis of metastatic disease alters the outlook for the patient so drastically, it should not be made without biopsy. Every oncologist has seen patients with tuberculosis, gallstones, primary hyperparathyroidism, or other nonmalignant diseases misdiagnosed and treated as though they had metastatic breast cancer. This is a catastrophic mistake and justifies biopsy for every patient at the time of initial suspicion of metastatic disease.

The choice of therapy requires consideration of local therapy needs, the overall medical condition of the patient, and the hormone receptor status of the tumor, as well as the exercise of clinical judgment. Because therapy of systemic disease is palliative, the potential toxicities of therapies should be balanced against the response rates. Several variables influence the response to systemic therapy. For example, the presence of estrogen and progesterone receptors is a strong indication for endocrine therapy, since the response rates for tumors that express both receptors may approach 70%. On the other hand, patients with short disease-free intervals, rapidly progressive visceral disease, lymphangitic pulmonary disease, or intracranial disease are unlikely to respond to endocrine therapy.

In many cases, systemic therapy can be withheld while the patient is managed with appropriate local therapy. Radiation therapy and occasionally surgery are effective at relieving the symptoms of metastatic disease, particularly when bony sites are involved. Many patients with bone-only or bone-dominant disease have a relatively indolent course. Under such circumstances, systemic chemotherapy has a modest effect, whereas radiation therapy may be effective for long periods. Other systemic treatments, such as strontium 89 and/or bisphosphonates, may provide a palliative benefit without inducing objective responses. Since the goal of therapy is to maintain well-being for as long as possible, emphasis should be placed on avoiding the most hazardous complications of metastatic disease, including pathologic fracture of the axial skeleton and spinal cord compression. New back pain in patients with cancer should be explored aggressively on an emergent basis; to wait for neurologic symptoms is a potentially catastrophic error. Metastatic involvement of endocrine organs can cause profound

dysfunction, including adrenal insufficiency and hypopituitarism. Similarly, obstruction of the biliary tree or other impaired organ function may be better managed with a local therapy than with a systemic approach.

Endocrine Therapy Normal breast tissue is estrogen-dependent. Both primary and metastatic breast cancer may retain this phenotype. The best means of ascertaining whether a breast cancer is hormone-dependent is through analysis of estrogen and progesterone receptor levels on the tumor. Tumors that are positive for the estrogen receptor and negative for the progesterone receptor have a response rate of approximately 30%. Tumors that have both receptors have a response rate approaching 70%. If neither receptor is present, the objective response rates are less than 10%. Receptor analyses provide information as to the correct ordering of endocrine therapies. Because of their lack of toxicity and because some patients whose receptor analyses are reported as negative respond to endocrine therapy, an endocrine treatment should be attempted in every patient with metastatic breast cancer. Potential endocrine therapies are summarized in Table 89-4. The choice of endocrine therapy is usually determined by toxicity profile and availability. In most patients, the initial endocrine therapy is the antiestrogen tamoxifen. Newer antiestrogens that are free of agonistic effects are in clinical trial. Cases in which tumors shrink in response to tamoxifen withdrawal (as well as withdrawal of pharmacologic doses of estrogens) have been reported. Endogenous estrogen formation may be blocked by aromatase inhibitors or analogues of luteinizing hormone-releasing hormone (LHRH). Additive endocrine therapies, including treatment with progestogens, estrogens, and androgens, may also be tried in patients who respond to initial endocrine therapy; the mechanism of action of these latter therapies is unknown. However, patients who respond to one endocrine therapy have at least a 50% chance of responding to a second endocrine therapy. It is not uncommon for patients to respond to two or three sequential endocrine therapies; however, combination endocrine therapies do not appear to be superior to individual agents, and combinations of chemotherapy with endocrine therapy are not useful. The median survival of patients with metastatic disease is approximately 2 years, and many patients, particularly older persons and those with hormone-dependent disease, may respond to endocrine therapy for 3 to 5 years or longer.

Chemotherapy Unlike many other epithelial malignancies, breast cancer responds to several chemotherapeutic agents, including anthracyclines, alkylating agents, taxanes, and antimetabolites. Multiple combinations of these agents have been found to improve response rates somewhat, but they have had little impact on duration of response or survival. As previously mentioned, median survival from diagnosis of metastatic disease is approximately 2 years. The choice among multidrug combinations frequently depends on whether adjuvant chemotherapy was administered and, if so, what type. While patients treated with adjuvant regimens such as cyclophosphamide, methotrexate, and fluorouracil (CMF regimens) may subsequently respond to the same combination in the metastatic disease setting, most oncologists use drugs to which the patients have not been previously exposed. Once patients have progressed after combination drug therapy, it is most common to treat them with single agents. Given the significant toxicity of most drugs, the use of a single effective agent will minimize toxicity by sparing the patient exposure to drugs that would be of little value. Unfortunately, no form of in vitro drug sensitivity testing to select the drugs most efficacious for a given patient has been demonstrated to be useful.

Most oncologists use either an anthracycline or paclitaxel following failure with the initial regimen. However, the choice has to be balanced with individual needs.

The use of a humanized antibody to *erbB2* (herceptin) combined with paclitaxel can improve response rate and survival for women whose metastatic tumors overexpress *erbB2*. The magnitude of the survival extension is modest in patients with metastatic disease. Application to adjuvant therapy may prove even more beneficial.

High-Dose Chemotherapy including Autologous Bone Marrow Transplantation
Autologous bone marrow transplantation combined with high doses of single agents can produce improvement even in heavily pretreated patients. However, such responses are rarely, if ever, durable and are unlikely to substantially alter the clinical course for most patients with advanced metastatic disease. Randomized trials have not been encouraging, and these approaches cannot be recommended as part of clinical care outside of research settings.

Stage III Breast Cancer Between 10 and 25% of patients have so-called locally advanced or stage III breast cancer at diagnosis. Many of these cancers are technically operable, whereas others, particularly cancers with chest wall involvement, inflammatory breast cancers, or cancers with large matted axillary lymph nodes, cannot be managed with surgery initially. Although no randomized trials have proved the efficacy of induction chemotherapy, this approach has gained widespread use. More than 90% of patients with locally advanced breast cancer show a partial or better response to multidrug chemotherapy regimens that include an anthracycline. Early administration of this treatment reduces the bulk of the disease and frequently makes the patient a suitable candidate for salvage surgery and/or radiation therapy. These patients should be managed in multimodality clinics, if possible, to coordinate surgery, radiation therapy, and systemic chemotherapy. Such approaches produce long-term disease-free survival in about 30 to 50% of patients.

Breast Cancer Prevention Women who have one breast cancer are at risk of developing a contralateral breast cancer at a rate of approximately 0.5% per year. When adjuvant tamoxifen is administered to these patients, the rate of development of contralateral breast cancers is reduced. In other tissues of the body, tamoxifen has estrogen-like effects that are beneficial: preservation of bone mineral density and long-term lowering of cholesterol. However, tamoxifen has estrogen-like effects on the uterus, leading to an increased risk of uterine cancer (0.75% incidence after 5 years on tamoxifen). The Breast Cancer Prevention Trial (BCPT) revealed a >40% reduction in breast cancer amongst women with a risk of at least 1.66% taking the drug for 5 years. Raloxifene has shown similar breast cancer prevention potency but may have different effects on bone and heart. The two are being compared in a prospective randomized prevention trial (the STAR trial).

Noninvasive Breast Cancer Breast cancer develops as a series of molecular changes in the epithelial cells that lead to ever more malignant behavior. Increased use of mammography and better mammographic diagnosis have led to more frequent diagnosis of noninvasive breast cancer. These lesions fall into two groups: ductal carcinoma in situ (DCIS) and lobular carcinoma in situ (lobular neoplasia). The

management of both entities is controversial.

Ductal Carcinoma in Situ Proliferation of cytologically malignant breast epithelial cells within the ducts is termedDCIS. Significant disagreement can occur in differentiating atypical hyperplasia from DCIS. At least one-third of the cases of untreated DCIS progress within 5 years to invasive breast cancer. For many years, the standard treatment for this disease was mastectomy. However, since treatment of this condition by lumpectomy and radiation therapy gives survival that is as good as the survival for invasive breast cancer by mastectomy, it appears paradoxical to recommend more aggressive therapy for a "less" malignant disease. In one randomized trial, the combination of wide excision plus irradiation for DCIS caused a substantial reduction in the local recurrence rate as compared with wide excision alone with negative margins, though survival is identical in the two arms. No studies have compared either of these regimens to mastectomy. Addition of tamoxifen to any DCIS surgical/radiation therapy regimen will further improve outcome.

Several prognostic features may help to identify patients at high risk for local recurrence after either lumpectomy alone or lumpectomy with radiation therapy. These include extensive disease; age less than 40; and cytologic features such as necrosis, poor nuclear grade, and comedo subtype with overexpression of erbB2. Some data suggest that adequate excision with careful determination of pathologically clear margins is associated with a low recurrence rates. When such surgery is combined with radiation therapy, recurrence (which is usually in the same quadrant) occurs with a frequency of £10%. Given the fact that half of these recurrences will be invasive, about 5% of the initial cohort will eventually develop invasive breast cancer. A reasonable expectation of mortality for these patients is about 1%, a figure that approximates the mortality rate for DCISmanaged by mastectomy. Although this train of reasoning has not formally been proved valid, it is reasonable at present to recommend that patients who desire breast preservation, and in whom DCIS appears to be reasonably localized, be managed by adequate surgery with meticulous pathologic evaluation, followed by breast irradiation and tamoxifen. For patients with localized DCIS, there is no need for axillary lymph node dissection. More controversial is the question of what management is optimal when there is any degree of invasion. Because of a significant likelihood (10 to 15%) of axillary lymph node involvement even when the primary lesion shows only microscopic invasion, it is prudent to do at least a level 1 and 2 axillary lymph node dissection for all patients with any degree of invasion, although in centers familiar with the technique, sentinel node biopsy may be substituted. Further management is dictated by the presence of nodal spread.

Lobular Neoplasia Proliferation of cytologically malignant cells within the lobules is termed *lobular neoplasia*. Approximately 30% of patients who have had adequate local excision of the lesion develop breast cancer (usually infiltrating ductal cell carcinoma) over the next 15 to 20 years. Ipsilateral and contralateral disease are equally common. Therefore, lobular neoplasia may be a premalignant lesion that suggests an elevated risk of subsequent breast cancer, rather than a form of malignancy itself, and aggressive local management seems unreasonable. Most patients should be treated with tamoxifen for 5 years and followed with careful annual mammography and semiannual physical examinations. Additional molecular analysis of these lesions may make it possible to discriminate between patients who are at risk of further progression

and who require additional therapy and those in whom simple follow-up is adequate.

Male Breast Cancer Breast cancer is about 1/150th as frequent in men as in women. It usually presents as a unilateral lump in the breast and is frequently not diagnosed promptly. Given the small amount of soft tissue and the unexpected nature of the problem, locally advanced presentations are somewhat more common. When male breast cancer is matched to female breast cancer by age and stage, its overall prognosis is identical. Although gynecomastia may initially be unilateral or asymmetric, any unilateral mass in a man over the age of 40 should receive a careful workup all the way through biopsy. On the other hand, bilateral symmetric breast development rarely represents breast cancer and is almost invariably due to endocrine disease or a drug effect. It should be kept in mind, nevertheless, that the risk of cancer is much greater in men with gynecomastia; in such men, gross asymmetry of the breasts should arouse suspicion of cancer. Male breast cancer is best managed by mastectomy and axillary lymph node dissection (modified radical mastectomy). Patients with locally advanced disease or positive nodes should also be treated with irradiation. Approximately 90% of male breast cancers contain estrogen receptors, and approximately 60% of cases with metastatic disease respond to endocrine therapy. There are no randomized studies exploring adjuvant therapy for male breast cancer. Two historic experiences suggest that the disease responds well to adjuvant systemic therapy, and, if not medically contraindicated, the same criteria for the use of adjuvant therapy in women should be applied to men.

The sites of relapse and spectrum of response to chemotherapeutic drugs are virtually identical for breast cancers in the two sexes.

FOLLOW-UP OF BREAST CANCER PATIENTS

Despite the availability of sophisticated and expensive imaging techniques and a wide range of serum tumor marker tests, no studies document that survival is influenced by early diagnosis of relapse. Surveillance guidelines are given in Table 89-5.

(Bibliography omitted in Palm version)

Back to Table of Contents

90. GASTROINTESTINAL TRACT CANCER - *Robert J. Mayer*

The gastrointestinal tract is the second most common noncutaneous site for cancer and the second major cause of cancer-related mortality in the United States.

ESOPHAGEAL CANCER

INCIDENCE AND ETIOLOGY

Cancer of the esophagus is a relatively uncommon but extremely lethal malignancy. The diagnosis was made in 12,300 Americans in 2000 and led to 12,100 deaths. Worldwide, the incidence of esophageal cancer varies strikingly. It occurs frequently within a geographic region extending from the southern shore of the Caspian Sea on the west to northern China on the east and encompassing parts of Iran, Central Asia, Afghanistan, Siberia, and Mongolia. High-incidence "pockets" of the disease are also present in such disparate locations as Finland, Iceland, Curacao, southeastern Africa, and northwestern France. In North America and western Europe, the disease is far more common in blacks than whites, is more common in males than females, appears most often after age 50, and seems to be associated with a lower socioeconomic status.

A variety of causative factors have been implicated in the development of the disease (Table 90-1). In the United States, esophageal cancer cases are either squamous cell carcinomas or adenocarcinomas. The etiology of squamous cell esophageal cancer is related to excess alcohol consumption and/or cigarette smoking. The relative risk increases with the amount of tobacco smoked or alcohol consumed, with these factors acting synergistically. The consumption of whiskey is linked to a higher incidence than the consumption of wine or beer. Squamous cell esophageal carcinoma has also been associated with the ingestion of nitrites, smoked opiates, and fungal toxins in pickled vegetables, as well as mucosal damage caused by such physical insults as long-term exposure to extremely hot tea, the ingestion of lye, radiation-induced strictures, and chronic achalasia. The presence of an esophageal web in association with glossitis and iron deficiency (i.e., Plummer-Vinson or Paterson-Kelly syndrome) and congenital hyperkeratosis and pitting of the palms and soles (i.e., tylosis palmaris et plantaris) have each been linked with squamous cell esophageal cancer, as have dietary deficiencies of molybdenum, zinc, and vitamin A.

For unclear reasons, the incidence of squamous cell esophageal cancer has decreased in both the black and white population in the United States over the past 20 years, while the rate of adenocarcinoma has risen dramatically, particularly in white males. Adenocarcinomas arise in the distal esophagus in the presence of chronic gastric reflux and gastric metaplasia of the epithelium (Barrett's esophagus), which is more common in obese persons. Adenocarcinomas arise within dysplastic columnar epithelium in the distal esophagus. Even before frank neoplasia is detectable, aneuploidy and p53 mutations are found in the dysplastic epithelium. These adenocarcinomas behave clinically like gastric adenocarcinoma and now account for >50% of esophageal cancers.

CLINICAL FEATURES

About 15% of esophageal cancers occur in the upper third of the esophagus (cervical esophagus), 40% in the middle third, and 45% in the lower third. Squamous cell carcinomas and adenocarcinomas of the esophagus cannot be distinguished radiographically or endoscopically.

Progressive dysphagia and weight loss of short duration are the initial symptoms in the vast majority of patients. Dysphagia initially occurs with solid foods and gradually progresses to include semisolids and liquids. By the time these symptoms develop, the disease is usually incurable, since difficulty in swallowing does not occur until ³60% of the esophageal circumference is infiltrated with cancer. Dysphagia may be associated with pain on swallowing (odynophagia), pain radiating to the chest and/or back, regurgitation or vomiting, and aspiration pneumonia. The disease most commonly spreads to adjacent and supraclavicular lymph nodes, liver, lungs, and pleura. Tracheoesophageal fistulas may develop as the disease advances, leading to severe suffering. As with other squamous cell carcinomas, hypercalcemia may occur in the absence of osseous metastases, probably from parathormone-related peptide secreted by tumor cells (Chap. 100).

DIAGNOSIS

Attempts at endoscopic and cytologic screening for carcinoma in patients with Barrett's esophagus, while effective as a means of detecting high-grade dysplasia, have not yet been shown to improve the prognosis in individuals found to have a carcinoma. Routine contrast radiographs effectively identify esophageal lesions large enough to cause symptoms. In contrast to benign esophageal leiomyomas, which result in esophageal narrowing with preservation of a normal mucosal pattern, esophageal carcinomas characteristically cause ragged, ulcerating changes in the mucosa in association with deeper infiltration, producing a picture resembling achalasia. Smaller, potentially resectable tumors are often poorly visualized despite technically adequate esophagograms. Because of this, esophagoscopy should be performed in all patients suspected of having an esophageal abnormality, to visualize the tumor and to obtain histopathologic confirmation of the diagnosis. Because the population of persons at risk for squamous cell carcinoma of the esophagus (i.e., smokers and drinkers) also has a high rate of cancers of the lung and the head and neck region, endoscopic inspection of the larynx, trachea, and bronchi should also be done. A thorough examination of the fundus of the stomach (by retroflexing the endoscope) is imperative as well. Endoscopic biopsies of esophageal tumors fail to recover malignant tissue in one-third of cases because the biopsy forceps cannot penetrate deeply enough through normal mucosa pushed in front of the carcinoma. Cytologic examination of tumor brushings frequently complements standard biopsies and should be performed routinely. The extent of tumor spread to the mediastinum and paraaortic lymph nodes should also be assessed by computed tomography (CT) scans of the chest and abdomen and by endoscopic ultrasound.

TREATMENT

The prognosis for patients with esophageal carcinoma is poor. Fewer than 5% of patients are alive 5 years after the diagnosis; thus, management focuses on symptom control. Surgical resection of all gross tumor (i.e., total resection) is feasible in only

40%of cases, with residual tumor cells frequently present at the resection margins. Such esophagectomies have been associated with a postoperative mortality rate of ~10% due to anastomotic fistulas, subphrenic abscesses, and respiratory complications. About 20% of patients who survive a total resection live 5 years. The outcome of primary radiation therapy (5500 to 6000 cGy) for squamous cell carcinomas is similar to that of radical surgery, sparing patients perioperative morbidity but often resulting in less satisfactory palliation of obstructive symptoms. The evaluation of chemotherapeutic agents in patients with esophageal carcinoma has been hampered by ambiguity in the definition of "response" (i.e., benefit) and the debilitated physical condition of many treated individuals. Nonetheless, significant reductions in the size of measurable tumor masses have been reported in 15 to 25% of patients given single-agent treatment and in 30 to 60% of patients treated with drug combinations that include cisplatin. Combination chemotherapy and radiation therapy as the initial therapeutic approach, either alone or followed by an attempt at operative resection, may be of benefit. When administered along with radiation therapy, chemotherapy produces a better survival outcome than radiation therapy alone. The use of preoperative chemotherapy and radiation therapy followed by esophageal resection appears to prolong survival as compared with historic controls, but randomized trials have produced inconsistent results.

For the incurable, surgically unresectable patient with esophageal cancer, dysphagia, malnutrition, and the management of tracheoesophageal fistulas loom as major issues. Approaches to palliation include repeated endoscopic dilatation, the surgical placement of a gastrostomy or jejunostomy for hydration and feeding, and endoscopic placement of an expansive metal stent to bypass the tumor. Endoscopic fulguration of the obstructing tumor with lasers appears to be the most promising of these techniques.

TUMORS OF THE STOMACH

GASTRIC ADENOCARCINOMA

Incidence and Epidemiology For unclear reasons, the incidence and mortality rates for gastric cancer have decreased markedly during the past 60 years. The mortality rate from gastric cancer in the United States has dropped in men from 28 to 5.0 per 100,000 population, while in women, the rate has decreased from 27 to 2.3 per 100,000. Nonetheless, 21,500 new cases of stomach cancer were diagnosed in the United States and 13,000 Americans died of the disease in 2000. Gastric cancer incidence has decreased worldwide but remains high in Japan, China, Chile, and Ireland.

The risk of gastric cancer is greater among lower socioeconomic classes. Migrants from high- to low-incidence nations maintain their susceptibility to gastric cancer, while the risk for their offspring approximates that of the new homeland. These findings suggest that an environmental exposure, probably beginning early in life, is related to the development of gastric cancer, with dietary carcinogens considered the most likely factor(s).

Pathology About 85% of stomach cancers are adenocarcinomas, with 15% due to lymphomas and leiomyosarcomas. Gastric adenocarcinomas may be subdivided into two categories: a *diffuse type* in which cell cohesion is absent, so that individual cells infiltrate and thicken the stomach wall without forming a discrete mass; and an *intestinal*

type characterized by cohesive neoplastic cells that form glandlike tubular structures. The diffuse carcinomas occur more often in younger patients, develop throughout the stomach (including the cardia), result in a loss of distensibility of the gastric wall (so-called linitis plastica or "leather bottle" appearance), and carry a poorer prognosis. Intestinal-type lesions are frequently ulcerative, more commonly appear in the antrum and lesser curvature of the stomach, and are often preceded by a prolonged precancerous process. While the incidence of diffuse carcinomas is similar in most populations, the intestinal type tends to predominate in the high-risk geographic regions and is less likely to be found in areas where the frequency of gastric cancer is declining. Thus, different etiologic factor(s) may be involved in these two subtypes. In the United States, the distal stomach is the site of origin of ~30% of gastric cancers, ~20% arise in the midportion of the stomach, and ~37% originate in the proximal third of the stomach. The remaining 13% involve the entire stomach.

Etiology The long-term ingestion of high concentrations of nitrates in dried, smoked, and salted foods appears to be associated with a higher risk. The nitrates are thought to be converted to carcinogenic nitrites by bacteria (Table 90-2). Such bacteria may be introduced exogenously through the ingestion of partially decayed foods, which are consumed in abundance worldwide by the lower socioeconomic classes. Bacteria such as *Helicobacter pylori* may also contribute to this effect by causing chronic gastritis, loss of gastric acidity, and bacterial growth in the stomach. Loss of acidity may occur when acid-producing cells of the gastric antrum have been removed surgically to control benign peptic ulcer disease or when achlorhydria, atrophic gastritis, and even pernicious anemia develop in the elderly. Serial endoscopic examinations of the stomach in patients with atrophic gastritis have documented replacement of the usual gastric mucosa by intestinal-type cells. This process of intestinal metaplasia may lead to cellular atypia and eventual neoplasia. Since the declining incidence of gastric cancer in the United States primarily reflects a decline in distal, ulcerating, intestinal-type lesions, it is conceivable that better food preservation and the availability of refrigeration to all socioeconomic classes have decreased the dietary ingestion of exogenous bacteria.

Several additional etiologic factors have been associated with gastric carcinoma. Gastric ulcers and adenomatous polyps have occasionally been so linked, but data regarding a cause-and-effect relationship are unconvincing. The inadequate clinical distinction between benign gastric ulcers and small ulcerating carcinomas may, in part, account for this presumed association. The presence of extreme hypertrophy of gastric rugal folds (i.e., Menetrier's disease), giving the impression of polypoid lesions, has been associated with a striking frequency of malignant transformation; such hypertrophy, however, does not represent the presence of true adenomatous polyps. Individuals with blood group A have a higher incidence of gastric cancer than persons with blood group O; this observation may be related to differences in the mucous secretion leading to altered mucosal protection from carcinogens. Duodenal ulcers are not associated with gastric cancer.

Clinical Features Gastric cancers, when superficial and surgically curable, usually produce no symptoms. As the tumor becomes more extensive, patients may complain of an insidious upper abdominal discomfort varying in intensity from a vague, postprandial fullness to a severe, steady pain. Anorexia, often with slight nausea, is very common but is not the usual presenting complaint. Weight loss may eventually be

observed, and nausea and vomiting are particularly prominent with tumors of the pylorus; dysphagia may be the major symptom caused by lesions of the cardia. There are no early physical signs. A palpable abdominal mass indicates long-standing growth and predicts regional extension.

Gastric carcinomas spread by direct extension through the gastric wall to the perigastric tissues, occasionally adhering to adjacent organs such as the pancreas, colon, or liver. The disease also spreads via lymphatics or by seeding of peritoneal surfaces. Metastases to intraabdominal and supraclavicular lymph nodes occur frequently, as do metastatic nodules to the ovary (Krukenberg's tumor), periumbilical region ("Sister Mary Joseph node") or peritoneal cul-de-sac (Blumer's shelf palpable on rectal or vaginal examination); malignant ascites may also develop. The liver is the most common site for hematogenous spread of tumor.

The presence of iron-deficiency anemia in men and of occult blood in the stool in both sexes mandate a search for an occult gastrointestinal tract lesion. A careful assessment is of particular importance in patients with atrophic gastritis or pernicious anemia. Unusual clinical features associated with gastric adenocarcinomas include migratory thrombophlebitis, microangiopathic hemolytic anemia, and acanthosis nigricans.

Diagnosis A double-contrast radiographic examination is the simplest diagnostic procedure for the evaluation of a patient with epigastric complaints. The use of double-contrast techniques helps to detect small lesions by improving mucosal detail. The stomach should be distended at some time during every radiographic examination, since decreased distensibility may be the only indication of a diffuse infiltrative carcinoma. Although gastric ulcers can be detected fairly early, distinguishing benign from malignant lesions is difficult. The anatomic location of an ulcer is not in itself an indication of the presence or absence of a cancer.

Gastric ulcers that appear benign by radiography present special problems. Some physicians believe that gastroscopy is not mandatory if the radiographic features are typically benign, if complete healing can be visualized by x-ray within 6 weeks, and if a follow-up contrast radiograph obtained several months later shows a normal appearance. However, we recommend gastroscopic biopsy and brush cytology for all patients with a gastric ulcer in order to exclude a malignancy. Malignant gastric ulcers must be recognized before they penetrate into surrounding tissues, because the rate of cure of early lesions limited to the mucosa or submucosa is>80%. Since gastric carcinomas are difficult to distinguish clinically or radiographically from gastric lymphomas, endoscopic biopsies should be made as deep as possible, due to the submucosal location of lymphoid tumors.

The staging system for gastric carcinoma is shown in Table 90-3.

TREATMENT

Complete surgical removal of the tumor with resection of adjacent lymph nodes offers the only chance for cure. However, this is possible in fewer than a third of patients. A subtotal gastrectomy is the treatment of choice for patients with distal carcinomas, while total or near-total gastrectomies are required for more proximal tumors. The inclusion of

extended lymph node dissection to these procedures appears to confer an added risk for complications without enhancing survival. The prognosis following complete surgical resection depends on the degree of tumor penetration into the stomach wall and is adversely influenced by regional lymph node involvement, vascular invasion, and abnormal DNA content (i.e., aneuploidy), characteristics found in the vast majority of American patients. As a result, the probability of survival after 5 years for the 25 to 30% of patients able to undergo complete resection is ~20% for distal tumors and<10% for proximal tumors, with recurrences continuing to occur for at least 8 years after surgery. In the absence of ascites or extensive hepatic or peritoneal metastases, however, even patients whose disease is believed to be incurable by surgery should be offered an attempt at resection of the primary lesion, since reduction of tumor bulk is the best form of palliation and may enhance the probability of benefit from chemotherapy and/or radiation therapy.

Gastric adenocarcinoma is a relatively radioresistant tumor, and adequate control of the primary tumor requires doses of external beam irradiation that exceed the tolerance of surrounding structures, such as bowel mucosa and spinal cord. As a result, the major role of radiation therapy in patients has been palliation of pain. Radiation therapy alone after a complete resection does not prolong survival. In the setting of surgically unresectable disease limited to the epigastrium, patients treated with 3500 to 4000 cGy did not live longer than similar patients not receiving radiotherapy; however, survival was prolonged slightly when 5-fluorouracil (5-FU) was given in combination with radiation therapy. In this clinical setting, the 5-FU may well be functioning as a radiosensitizer.

The administration of combinations of cytotoxic drugs to patients with advanced gastric carcinoma has been associated with partial responses in 30 to 50% of cases, providing significant benefit to individuals who respond to treatment. Such drug combinations have generally included5-FU and doxorubicin together with mitomycin-C, cisplatin, or high doses of methotrexate. Despite this encouraging response rate, complete remissions are uncommon, the partial responses are transient, and the overall influence of multidrug therapy on survival has been a source of debate. The use of prophylactic (i.e., adjuvant) chemotherapy following the complete resection of a gastric cancer has not improved survival. However, postoperative chemotherapy combined with radiation therapy has been shown to reduce the recurrence rate and prolong survival.

PRIMARY GASTRIC LYMPHOMA

Primary lymphoma of the stomach is relatively uncommon, accounting for<15% of gastric malignancies and about 2% of all lymphomas. The stomach is, however, the most frequent extranodal site for lymphoma, and gastric lymphoma has increased in frequency during the past 25 years. The disease is difficult to distinguish clinically from gastric adenocarcinoma; both tumors are most often detected during the sixth decade of life; present with epigastric pain, early satiety, and generalized fatigue; and are usually characterized by ulcerations with a ragged, thickened mucosal pattern demonstrated by contrast radiographs. The diagnosis of lymphoma of the stomach may occasionally be made through cytologic brushings of the gastric mucosa but usually it requires a biopsy at gastroscopy or laparotomy. Failure of gastroscopic biopsies to detect lymphoma in a given case should not be interpreted as being conclusive, since superficial biopsies may

miss the deeper lymphoid infiltrate. The macroscopic pathology of gastric lymphoma may also mimic adenocarcinoma, consisting of either a bulky ulcerated lesion localized in the corpus or antrum or a diffuse process spreading throughout the entire gastric submucosa and even extending into the duodenum. Microscopically, the vast majority of gastric lymphoid tumors are non-Hodgkin's lymphomas of B cell origin; Hodgkin's disease involving the stomach is extremely uncommon. Histologically, these tumors may range from well-differentiated, superficial processes [mucosa-associated lymphoid tissue (MALT)] to high-grade, large cell lymphomas. Infection with *H. pylori*, the same bacterium associated with the development of gastric adenocarcinoma, appears to increase the risk for gastric lymphoma in general andMALTlymphomas in particular. Gastric lymphomas spread initially to regional lymph nodes (often to Waldeyer's ring) and may then disseminate. Gastric lymphomas are staged like other lymphomas (Chap. 112).

TREATMENT

Primary gastric lymphoma is a far more treatable disease than adenocarcinoma of the stomach, a fact that underscores the need for making the correct diagnosis. Antibiotic treatment to eradicate *H. pylori* infection has led to regression of about 75% of gastricMALTlymphomas and should be considered before surgery, radiation therapy, or chemotherapy are undertaken in patients having such tumors. Responding patients should undergo periodic endoscopic surveillance because it remains unclear whether the neoplastic clone is eliminated or merely suppressed. Subtotal gastrectomy, usually followed by combination chemotherapy, has led to 5-year survival rates of 40 to 60% in patients with localized high-grade lymphomas. The need for a major surgical procedure is not clear, particularly in patients with preoperative radiographic evidence of nodal involvement, for whom chemotherapy alone is effective therapy. A role for radiation therapy is not defined because most recurrences develop at sites distant from the epigastrium. If widespread disease is discovered at the time of laparotomy, combination chemotherapy should be used.

GASTRIC (NONLYMPHOID) SARCOMA

Leiomyosarcomas are the most common of this group of gastric malignancies and make up 1 to 3% of gastric neoplasms. They most frequently involve the anterior and posterior walls of the gastric fundus and often ulcerate and bleed. Even those lesions that appear benign on histologic examination may behave in a malignant fashion. Leiomyosarcomas rarely invade adjacent viscera and characteristically do not metastasize to lymph nodes, but they may spread to the liver and lungs. The treatment of choice is surgical resection. Combination chemotherapy should be reserved for patients with metastatic disease.

COLORECTAL CANCER

INCIDENCE

Cancer of the large bowel is second only to lung cancer as a cause of cancer death in the United States. Approximately 130,200 new cases occurred in 2000, and 56,300 deaths were due to colorectal cancer. The incidence rate has declined slightly during the past 15 years and the mortality rate has decreased in recent years, particularly in

females. Colorectal cancer generally occurs in individuals³50 years.

POLYPS AND MOLECULAR PATHOGENESIS

Most colorectal cancers, regardless of etiology, arise from adenomatous polyps. A polyp is a grossly visible protrusion from the mucosal surface and may be classified pathologically as a nonneoplastic hamartoma (*juvenile polyp*), a hyperplastic mucosal proliferation (*hyperplastic polyp*), or an adenomatous polyp. Only adenomas are clearly premalignant, and only a minority of such lesions ever develop into cancer. Population-screening studies and autopsy surveys have revealed that adenomatous polyps may be found in the colons of>30% of middle-aged or elderly people; however<1% of polyps ever become malignant. Most polyps produce no symptoms and remain clinically undetected. Occult blood in the stool may be found in<5% of patients with such lesions.

A number of molecular changes have been described in DNA obtained from adenomatous polyps, dysplastic lesions, and polyps containing microscopic foci of tumor cells (carcinoma in situ), which are thought to represent a multistep process in the evolution of normal colonic mucosa to life-threatening invasive carcinoma. These developmental steps towards carcinogenesis include point mutations in the K-*ras* protooncogene; hypomethylation of DNA, leading to gene activation; loss of DNA ("allelic loss") at the site of a tumor suppressor gene [the adenomatous polyposis coli (*APC*) gene] located on the long arm of chromosome 5 (5q21); allelic loss at the site of a tumor suppressor gene located on chromosome 18q [the deleted in colorectal cancer (*DCC*) gene]; and allelic loss at chromosome 17p, associated with mutations in the *p53* tumor suppressor gene (Chap. 81). Thus, the altered proliferative pattern of the colonic mucosa, which results in progression to a polyp and then to carcinoma, may involve the mutational activation of an oncogene followed by and coupled with the loss of genes that normally suppress tumorigenesis. While the present model includes five such molecular alterations, others are likely involved in the carcinogenic process. It remains uncertain whether the genetic aberrations always occur in a defined order. Based on this model, however, it is believed that neoplasia develops only in those polyps in which all of these mutational events take place.

Clinically, the probability of an adenomatous polyp becoming a cancer depends on the gross appearance of the lesion, its histologic features, and its size. Adenomatous polyps may be pedunculated (stalked) or sessile (flat-based). Cancers develop more frequently in sessile polyps. Histologically, adenomatous polyps may be tubular, villous (i.e., papillary), or tubulovillous. Villous adenomas, most of which are sessile, become malignant more than three times as often as tubular adenomas. The likelihood that any polypoid lesion in the large bowel contains invasive cancer is related to the size of the polyp, being negligible (<2%) in lesions<1.5 cm, intermediate (2 to 10%) in lesions 1.5 to 2.5 cm in size, and substantial (10%) in lesions >2.5 cm.

Following the detection of an adenomatous polyp, the entire large bowel should be visualized endoscopically or radiographically, since synchronous lesions are present in about one-third of cases. Colonoscopy should then be repeated periodically, even in the absence of a previously documented malignancy, since such patients have a 30 to 50% probability of developing another adenoma and are at a higher-than-average risk for

developing a colorectal carcinoma. Adenomatous polyps are thought to require >5 years of growth before becoming clinically significant; colonoscopy need not be carried out more frequently than every 3 years.

ETIOLOGY AND RISK FACTORS

Risk factors for the development of colorectal cancer are listed in Table 90-4.

Diet The etiology for most cases of large-bowel cancer appears to be related to environmental factors. The disease occurs more often in upper socioeconomic populations who live in urban areas. Mortality from colorectal cancer is directly correlated with per capita consumption of calories, meat protein, and dietary fat and oil as well as elevations in the serum cholesterol concentration and mortality from coronary artery disease. Geographic variations in incidence are unrelated to genetic differences, since migrant groups tend to assume the large-bowel cancer incidence rates of their adopted countries. Furthermore, population groups such as Mormons and Seventh Day Adventists, whose lifestyle and dietary habits differ somewhat from those of their neighbors, have significantly lower than expected incidence and mortality rates for colorectal cancer. Colorectal cancer has increased in Japan since that nation has adopted a more "western" diet. At least two hypotheses have been proposed to explain the relationship to diet, neither of which is fully satisfactory.

Animal Fats One hypothesis is that the ingestion of animal fats leads to an increased proportion of anaerobes in the gut microflora, resulting in the conversion of normal bile acids into carcinogens. This provocative hypothesis is supported by several reports of increased amounts of fecal anaerobes in the stools of patients with colorectal cancer. Diets high in animal (but not vegetable) fats are also associated with high serum cholesterol, which is also associated with enhanced risk for the development of colorectal adenomas and carcinomas.

Fiber The observation that South African Bantus ingest a diet far higher in roughage, produce more frequent, bulkier stools, and have a lower incidence of large-bowel cancer than Americans and Europeans led to the proposal that the higher rate of colorectal cancer in western society results from low intake of dietary fiber. This theory suggests that dietary fiber accelerates intestinal transit time, thereby reducing the exposure of colonic mucosa to potential carcinogens and diluting these carcinogens because of enhanced fecal bulk. This theory has been largely discredited. Although an enhanced fiber intake increases fecal bulk, higher fiber intake has not been documented to consistently shorten stool transit time. In addition, despite the generally higher fiber intake in low-incidence countries, the environmental differences between developing and industrialized nations are myriad and include such other important dietary variables as meat and fat consumption. Furthermore, a diet low in fiber may lead to chronic constipation and diverticulosis. If a low-fiber diet were a significant risk factor in colorectal cancer, individuals with diverticulosis should be at higher risk for developing colorectal tumors; this is not the case. Finally, addition of fiber to the diet does not protect against the development of adenomatous polyps or colorectal cancer.

Thus, the weight of epidemiologic evidence implicates diet as being the major etiologic factor for colorectal cancer, particularly diets high in calories and animal fat.

HEREDITARY FACTORS AND SYNDROMES

As many as 25% of patients with colorectal cancer have a family history of the disease, suggesting a hereditary predisposition. Inherited large-bowel cancers can be divided into two main groups: the well-studied but uncommon polyposis syndromes and the more common nonpolyposis syndromes (Table 90-5).

Polyposis Coli Polyposis coli (familial polyposis of the colon) is a rare condition characterized by the appearance of thousands of adenomatous polyps throughout the large bowel. It is transmitted as an autosomal dominant trait; the occasional patients with no family history probably developed the condition due to a spontaneous mutation. Polyposis coli is associated with a deletion in the long arm of chromosome 5 (including the *APC* gene) in both neoplastic (somatic mutation) and normal (germline mutation) cells. The loss of this genetic material (i.e., allelic loss) results in the absence of tumor suppressor genes whose protein products would normally inhibit neoplastic growth. The presence of soft tissue and bony tumors, congenital hypertrophy of the retinal pigment epithelium, mesenteric desmoid tumors, and of ampullary cancers in addition to the colonic polyps characterizes a subset of polyposis coli known as *Gardner's syndrome*. The appearance of malignant tumors of the central nervous system accompanying polyposis coli defines *Turcot's syndrome*. The colonic polyps in all these conditions are rarely present before puberty but are generally evident in affected individuals by age 25. If the polyposis is not treated surgically, colorectal cancer will develop in almost all patients before age 40. Polyposis coli results from a defect in the colonic mucosa leading to an abnormal proliferative pattern and an impaired DNA repair following exposure to radiation or ultraviolet light. Once the multiple polyps that constitute polyposis coli are detected, patients should undergo a total colectomy. The ileoanal anastomotic technique allows removal of the entire bowel while retaining the anal sphincter; this appears to be the best treatment. Medical therapy with nonsteroidal anti-inflammatory drugs such as sulindac and cyclooxygenase-2 inhibitors such as celecoxib decreases the number and size of polyps in patients with polyposis coli; however, this effect on polyps is only temporary. Colectomy remains the primary therapy. The offspring of patients with polyposis coli, who often are prepubertal when the diagnosis is made in the parent, have a 50% risk for the development of this premalignant disorder and should be carefully screened by annual flexible sigmoidoscopy until age 35. Proctosigmoidoscopy is a sufficient screening procedure because polyps tend to be evenly distributed from cecum to anus, making more invasive and expensive techniques such as colonoscopy or barium enema unnecessary. Testing for occult blood in the stool is an inadequate screening maneuver. An alternative method for identifying carriers is testing DNA from peripheral blood mononuclear cells for the presence of a mutated *APC* gene. The detection of such a germline mutation can lead to a definitive diagnosis before the development of polyps.

Hereditary Nonpolyposis Colon Cancer Hereditary nonpolyposis colon cancer (HNPCC), also known as Lynch syndrome, is another autosomal dominant trait. It is characterized by the presence of three or more relatives with histologically documented colorectal cancer, one of whom is a first-degree relative of the other two; one or more cases of colorectal cancer diagnosed before age 50 in the family; and colorectal cancer involving at least two generations. In contrast to polyposis coli, HNPCC is associated

with an unusually high frequency of cancer arising in the proximal large bowel. The median age for the appearance of an adenocarcinoma is<50 years, 10 to 15 years younger than the median age for the general population. Despite having a poorly differentiated histologic appearance, the proximal colon tumors in HNPCC have a better prognosis than sporadic tumors from patients of similar age. Families with HNPCC often include individuals with multiple primary cancers; the association of colorectal cancer with either ovarian or endometrial carcinomas is especially strong in women. It has been recommended that members of such families undergo biennial colonoscopy beginning at age 25 years, with intermittent pelvic ultrasonography and endometrial biopsy offered for potentially afflicted women; such a screening strategy has not yet been validated. HNPCC is associated with germline mutations of several genes, particularly *hMSH2* on chromosome 2 and *hMLH1* on chromosome 3. These mutations lead to errors in DNA replication and are thought to result in DNA instability because of defective repair of DNA mismatches, resulting in abnormal cell growth and tumor development. Testing tumor cells for "microsatellite instability" (sequence changes reflecting defective mismatch repair) in patients under age 50 with colorectal cancer and a positive family history for colorectal or endometrial cancer may identify probands with HNPCC.

INFLAMMATORY BOWEL DISEASE (See alsoChap. 287)

Large-bowel cancer is increased in incidence in patients with long-standing inflammatory bowel disease. Cancers develop more commonly in patients with ulcerative colitis than in those with granulomatous colitis, but this impression may result in part from the occasional difficulty of differentiating these two conditions. The risk of colorectal cancer in a patient with inflammatory bowel disease is relatively small during the initial 10 years of the disease, but then it appears to increase at a rate of ~0.5 to 1% per year. Cancer may develop in 8 to 30% of patients after 25 years. The risk is higher in younger patients with pancolitis.

Cancer surveillance in patients with inflammatory bowel disease is unsatisfactory. Symptoms such as bloody diarrhea, abdominal cramping, and obstruction, which may signal the appearance of a tumor, are similar to the complaints caused by a flare-up of the underlying disease. In patients with a history of inflammatory bowel disease lasting 15 years or more who continue to experience exacerbations, the surgical removal of the colon can significantly reduce the risk for cancer and also eliminate the target organ for the underlying chronic gastrointestinal disorder. The value of such surveillance techniques as colonoscopy with mucosal biopsies and brushings for less symptomatic individuals with chronic inflammatory bowel disease is uncertain. The lack of uniformity regarding the pathologic criteria that characterize dysplasia and the absence of data that such surveillance reduces the development of lethal cancers have made this costly practice an area of controversy.

OTHER HIGH-RISK CONDITIONS

***Streptococcus bovis* Bacteremia** For unknown reasons, individuals who develop endocarditis or septicemia from this fecal bacteria have a high incidence of occult colorectal tumors and, possibly, upper gastrointestinal cancers as well. Endoscopic or radiographic screening appears advisable.

Ureterosigmoidostomy There is a 5 to 10% incidence of colon cancer 15 to 30 years after ureterosigmoidostomy to correct congenital extrophy of the bladder. Neoplasms characteristically are found at a site distal to the ureteral implant where colonic mucosa is chronically exposed to both urine and feces.

Tobacco Use Cigarette smoking is linked to the development of colorectal adenomas, particularly after more than 35 years of tobacco use. No biologic explanation for this association has yet been proposed.

PRIMARY PREVENTION

Several orally administered compounds have been assessed as possible inhibitors of colon cancer. The most effective class of these chemopreventive agents is aspirin and other nonsteroidal anti-inflammatory drugs, which are thought to suppress cell proliferation by inhibiting prostaglandin synthesis. Regular aspirin use reduces the risk for colonic adenomas and carcinomas as well as for death from large-bowel cancer; this inhibiting effect on colonic carcinogenesis appears to increase with the duration of drug use. Oral folic acid supplements and oral calcium supplements have been found to reduce the risk of adenomatous polyps and colorectal cancers in case-control studies. While antioxidant vitamins such as ascorbic acid, tocopherols, and b-carotene are present in diets rich in fruits and vegetables, which have been associated with lower rates of colorectal cancer, they have been found to be ineffective at reducing the incidence of subsequent adenomas in patients who had undergone the removal of a colonic adenoma. Estrogen replacement therapy has been associated with a reduction in the risk of colorectal cancer in women, conceivably by an effect on bile acid synthesis and composition. The otherwise unexplained reduction in colorectal cancer mortality in women may be a result of the widespread use of estrogen replacement in postmenopausal individuals.

SCREENING

The rationale for colorectal cancer screening programs is that the earlier detection of localized, superficial cancers in asymptomatic individuals will increase the surgical cure rate. Such screening programs are important for individuals having a family history of the disease in first-degree relatives. The relative risk for developing colorectal cancer increases to 1.75 in such people and may be even higher if the relative was afflicted before age 60. The use of proctosigmoidoscopy as a screening tool was based on the observation that 60% of early lesions are located in the rectosigmoid. For unexplained reasons, however, the proportion of large-bowel cancers arising in the rectum has been decreasing during the past several decades, with a corresponding increase in the proportion of cancers in the more proximal descending colon. As such, the potential for rigid proctosigmoidoscopy to detect a sufficient number of occult neoplasms to make the procedure cost-effective has been questioned. Flexible, fiberoptic sigmoidoscopes permit trained operators to visualize the colon for up to 60 cm, which enhances the capability for cancer detection. However, this technique still leaves the proximal half of the large bowel unscreened.

Most programs directed at the early detection of colorectal cancers have focused on digital rectal examinations and fecal occult blood testing. The digital examination should

be part of any routine physical evaluation in adults older than age 40, serving as a screening test for prostate cancer in men, a component of the pelvic examination in women, and an inexpensive maneuver for the detection of masses in the rectum. The development of the Hemoccult test has greatly facilitated the detection of occult fecal blood. Unfortunately, even when performed optimally, the Hemoccult test has major limitations as a screening technique. About 50% of patients with documented colorectal cancers have a negative fecal Hemoccult test, consistent with the intermittent bleeding pattern of these tumors. When random cohorts of asymptomatic persons have been tested, 2 to 4% have Hemoccult-positive stools. Colorectal cancers have been found in<10% of these "test-positive" cases, with benign polyps being detected in an additional 20 to 30%. Thus, a colorectal neoplasm will not be found in most asymptomatic individuals with occult blood in their stool. Nonetheless, persons found to have Hemoccult-positive stool routinely undergo further medical evaluation, including sigmoidoscopy, barium enema, and/or colonoscopy -- procedures that are not only uncomfortable and expensive but also associated with a small risk for significant complications. The added cost of these studies would appear justifiable if the small number of patients found to have occult neoplasms because of Hemoccult screening could be shown to have an improved prognosis and prolonged survival. Prospectively controlled trials addressing this issue have been performed. One of these studies, conducted at the University of Minnesota and involving >46,000 participants, reported a statistically significant reduction in mortality from colorectal cancer for individuals undergoing annual screening. However, this benefit only emerged after >13 years of follow-up and was extremely expensive to achieve, since all positive tests (most of which were false-positive) were followed by colonoscopy. Moreover, these colonoscopic examinations may have represented "chance selection" for more effective endoscopic screening and may also have provided the opportunity for cancer prevention through the removal of potentially premalignant adenomatous polyps.

Screening techniques for large-bowel cancer in asymptomatic persons remain unsatisfactory. Recommendations from governmental and private agencies are conflicting. Compliance with any screening strategy within the general population is poor. At present, the American Cancer Society suggests annual digital rectal examinations beginning at age 40, annual fecal Hemoccult screening beginning at age 50, and sigmoidoscopy (preferably flexible) every 3 to 5 years beginning at age 50 for asymptomatic individuals having no colorectal cancer risk factors. The use of colonscopy or double-contrast barium enemas for screening have not yet been systematically examined. Nonetheless, the American Cancer Society has proposed such a "total colon examination" every 10 years as an alternative to Hemoccult testing with periodic flexible sigmoidoscopy. More effective techniques for screening are needed, perhaps taking advantage of the molecular changes that have been described in these tumors. Analysis of stool for specific *ras* protooncogene mutations is being tested.

CLINICAL FEATURES

Presenting Symptoms Symptoms vary with the anatomic location of the tumor. Since stool is relatively liquid as it passes through the ileocecal valve into the right colon, cancers arising in the cecum and ascending colon may become quite large, without resulting in any obstructive symptoms or noticeable alterations in bowel habits. Lesions

of the right colon commonly ulcerate, leading to chronic, insidious blood loss without a change in the appearance of the stool. Consequently, patients with tumors of the ascending colon often present with symptoms such as fatigue, palpitations, and even angina pectoris and are found to have a hypochromic, microcytic anemia indicative of iron deficiency. Since the cancer may bleed intermittently, a random fecal occult blood test may be negative. As a result, the unexplained presence of iron-deficiency anemia in any adult (with the possible exception of a premenopausal, multiparous woman) mandates a thorough endoscopic and/or radiographic visualization of the entire large bowel (Fig. 90-1).

Since stool becomes more concentrated as it passes into the transverse and descending colon, tumors arising there tend to impede the passage of stool, resulting in the development of abdominal cramping, occasional obstruction, and even perforation. Radiographs of the abdomen often reveal characteristic annular, constricting lesions ("apple-core" or "napkin-ring") (Fig. 90-2).

Cancers arising in the rectosigmoid are often associated with hematochezia, tenesmus, and narrowing of the caliber of stool; anemia is an infrequent finding. While these symptoms may lead patients and their physicians to suspect the presence of hemorrhoids, the development of rectal bleeding and/or altered bowel habits demands a prompt digital rectal examination and proctosigmoidoscopy.

Staging, Prognostic Factors, and Patterns of Spread The prognosis for individuals having colorectal cancer is related to the depth of tumor penetration into the bowel wall and the presence of both regional lymph node involvement and distant metastases. These variables are incorporated into the staging system introduced by Dukes and applied to a TNM classification method, in which T represents the depth of tumor penetration, N the presence of lymph node involvement, and M the presence or absence of distant metastases (Table 90-6). Superficial lesions that do not penetrate into the muscularis or involve regional lymph nodes are designated as *stage A* (T1N0M0) disease; tumors that penetrate more deeply but have not spread to lymph nodes are *stage B* disease [subclassified as *stage* B_1(T2N0M0) if lesions are restricted to the muscularis and as *stage* B_2(T3N0M0) if lesions involve or penetrate the serosa]; regional lymph node involvement defines *stage C* (TxN1M0) disease; and metastatic spread to sites such as liver, lung, or bone indicates *stage D* (TxNxM1) disease. Unless gross evidence of metastatic disease is present, disease stage cannot be determined accurately before surgical resection and pathologic analysis of the operative specimens. It is not clear whether the detection of nodal metastases by special immunohistochemical molecular techniques has the same prognostic implications as disease detected by routine light microscopy.

Most recurrences after a surgical resection of a large-bowel cancer occur within the first 4 years, making 5-year survival a fairly reliable indicator of cure. The likelihood for 5-year survival in patients with colorectal cancer is stage-related (Table 90-6). That likelihood has improved during the past several decades when similar surgical stages have been compared. The most plausible explanation for this improvement appears to be more thorough intraoperative and pathologic staging. In particular, more exacting attention to pathologic detail has revealed that the prognosis following the resection of a colorectal cancer is not related merely to the presence or absence of regional lymph

node involvement but may be more precisely assessed by the number of involved lymph nodes (one to four lymph nodes versus five or more lymph nodes). Other predictors of a poor prognosis after a total surgical resection include tumor penetration through the bowel wall into pericolic fat, poorly differentiated histology, perforation and/or tumor adherence to adjacent organs (increasing the risk for an anatomically adjacent recurrence), and venous invasion by tumor (Table 90-7). Regardless of the clinicopathologic stage, a preoperative elevation of the plasma carcinoembryonic antigen (CEA) level predicts eventual tumor recurrence. The presence of aneuploidy and specific chromosomal deletions, such as allelic loss in chromosome 18q (involving the *DCC* gene) in tumor cells, appears to predict a higher risk for metastatic spread, particularly in patients with stage B_2(T3N0M0) disease. Conversely, the detection of microsatellite instability in tumor tissue has been associated with a more favorable outcome. In contrast to most other cancers, the prognosis in colorectal cancer is not influenced by the size of the primary lesion when adjusted for nodal involvement and histologic differentiation.

Cancers of the large bowel generally spread to regional lymph nodes or to the liver via the portal venous circulation. The liver represents the most frequent visceral site of metastatic dissemination; it is the initial site of distant spread in one-third of recurring colorectal cancers and is involved in more than two-thirds of such patients at the time of death. In general, colorectal cancer rarely metastasizes to the lungs, supraclavicular lymph nodes, bone, or brain without prior spread to the liver. A major exception to this rule occurs in patients having primary tumors in the distal rectum, from which tumor cells may spread through the paravertebral venous plexus, escaping the portal venous system and thereby reaching the lungs or supraclavicular lymph nodes without hepatic involvement. The median survival after the detection of distant metastases is 6 to 9 months (hepatomegaly, abnormal liver chemistries) to 24 to 30 months (small liver nodule initially identified by elevatedCEAlevel and subsequentCTscan).

TREATMENT

Total resection of tumor is the optimal treatment when a malignant lesion is detected endoscopically or radiographically in the large bowel. An evaluation for the presence of metastatic disease, including a thorough physical examination, chest radiograph, biochemical assessment of liver function, and measurement of the plasmaCEAlevel, should be performed before surgery. When possible, a colonoscopy of the entire large bowel should be performed to identify synchronous neoplasms and/or polyps. The detection of metastases should not preclude surgery in patients with tumor-related symptoms such as gastrointestinal bleeding or obstruction, but it often prompts the use of a less radical operative procedure. At the time of laparotomy, the entire peritoneal cavity should be examined, with thorough inspection of the liver, pelvis, and hemidiaphragm and careful palpation of the full length of the large bowel. Following recovery from a complete resection, patients should be observed carefully for 5 years by semiannual physical examinations and yearly blood chemistry measurements. If a complete colonoscopy was not performed preoperatively, it should be carried out within the first several postoperative months. Some authorities favor measuring plasma CEA levels at 3-month intervals because of the sensitivity of this test as a marker for otherwise undetectable tumor recurrence. Subsequent endoscopic or radiographic surveillance of the large bowel, probably at triennial intervals, is indicated, since patients

who have been cured of one colorectal cancer have a 3 to 5% probability of developing an additional bowel cancer during their lifetime and a>15% risk for the development of adenomatous polyps. Anastomotic ("suture-line") recurrences are infrequent in colorectal cancer patients provided the surgical resection margins were adequate and free of tumor. PeriodicCTscreening, chest radiographs, or more frequent colonoscopic examinations do not affect prognosis and add unnecessary costs to postoperative surveillance.

Radiation therapy to the pelvis is recommended for patients with rectal cancer because it reduces the 30 to 40% probability of regional recurrences following complete surgical resection of stage B or C tumors, especially if they have penetrated through the serosa. This alarmingly high rate of local disease recurrence is believed to be due to the fact that the contained anatomic space within the pelvis limits the extent of the resection and because the rich lymphatic network of the pelvic side wall immediately adjacent to the rectum facilitates the early spread of malignant cells into surgically inaccessible tissue. Radiation therapy, either pre- or postoperatively, reduces the likelihood of pelvic recurrences but does not appear to prolong survival. Preoperative radiotherapy is indicated for patients with large, potentially unresectable rectal cancers; such lesions may shrink enough to permit subsequent surgical removal. Radiation therapy is not effective in the primary treatment of colon cancer.

Chemotherapy in patients with advanced colorectal cancer has proven to be of only marginal benefit.5-FU is the most effective single agent for this disease. Partial responses are obtained in 15 to 20% of patients. The probability of tumor response appears to be somewhat greater for patients with liver metastases when chemotherapy is infused directly into the hepatic artery, but intraarterial treatment is costly and toxic and does not appear to prolong survival. The concomitant administration of folinic acid (leucovorin) improves the efficacy of 5-FU in patients with advanced colorectal cancer, presumably by enhancing the binding of 5-FU to its target enzyme, thymidylate synthase. A threefold improvement in the partial response rate is noted when folinic acid is combined with 5-FU; however, the effect on survival is marginal, and the optimal dose schedule remains to be defined.

Irinotecan (CPT-11), a topoisomerase 1 inhibitor, prolongs survival when compared to supportive care in patients whose disease has progressed on5-FU. Furthermore, the addition of irinotecan to 5-FU and leucovorin improves response rates and survival of patients with metastatic disease. Oxaliplatin, a platinum analogue, also improves the response rate when added to 5-FU and leucovorin as initial treatment of patients with metastatic disease.

Patients with solitary hepatic metastases without clinical or radiographic evidence of additional tumor involvement should be considered for partial liver resection, because such procedures are associated with 5-year survival rates of 25 to 30% when performed on selected individuals by experienced surgeons.

The administration of5-FU and leucovorin for 6 months after resection of tumor in patients with stage C disease leads to a 40% decrease in recurrence rates and 30% improvement in survival. Patients with stage B_2tumors do not benefit from adjuvant therapy. In rectal cancer, the delivery of postoperative (and probably preoperative)

combined modality therapy (5-FU plus radiation therapy) reduces the risk of recurrence and increases the chance of cure for patients with stages B_2 and C tumors. The 5-FU acts as a radiosensitizer when delivered together with radiation therapy.

TUMORS OF THE SMALL INTESTINE

Small-bowel tumors comprise <5% of gastrointestinal neoplasms. Because of their rarity, a correct diagnosis is often delayed. Abdominal symptoms are usually vague and poorly defined, and conventional radiographic studies of the upper and lower intestinal tract often appear normal. Small-bowel tumors should be considered in the differential diagnosis in the following situations: (1) recurrent, unexplained episodes of crampy abdominal pain; (2) intermittent bouts of intestinal obstruction, especially in the absence of inflammatory bowel disease or prior abdominal surgery; (3) intussusception in the adult; and (4) evidence of chronic intestinal bleeding in the presence of negative conventional contrast radiographs. A careful small-bowel barium study is the diagnostic procedure of choice; the diagnostic accuracy may be improved by infusing barium through a nasogastric tube placed into the duodenum (enteroclysis).

BENIGN TUMORS

The histology of benign small-bowel tumors is difficult to predict on clinical and radiologic grounds alone. The symptomatology of benign tumors is not distinctive, with pain, obstruction, and hemorrhage being the most frequent symptoms. These tumors are usually discovered during the fifth and sixth decades of life, more often in the distal rather than the proximal small intestine. The most common benign tumors are adenomas, leiomyomas, lipomas, and angiomas.

Adenomas These tumors include those of the islet cells and Brunner's glands as well as polypoid adenomas. *Islet cell adenomas* are occasionally located outside the pancreas; the associated syndromes are discussed inChap. 93. *Brunner's gland adenomas* are not truly neoplastic but represent a hypertrophy or hyperplasia of submucosal duodenal glands. These appear as small nodules in the duodenal mucosa that secrete a highly viscous alkaline mucus. Most often, this is an incidental radiographic finding not associated with any specific clinical disorder.

Polypoid Adenomas About 25% of benign small-bowel tumors are polypoid adenomas (Table 90-5). They may present as single polypoid lesions or, less commonly, as papillary villous adenomas. As in the colon, the sessile or papillary form of the tumor is sometimes associated with a coexisting carcinoma. Occasionally, patients with Gardner's syndrome develop premalignant adenomas in the small bowel; such lesions are generally in the duodenum. Multiple polypoid tumors may occur throughout the small bowel (and occasionally the stomach and colorectum) in the Peutz-Jeghers syndrome. The polyps are usually hamartomas (juvenile polyps) having a low potential for malignant degeneration. Mucocutaneous melanin deposits as well as tumors of the ovary, breast, pancreas, and endometrium are also associated with this autosomal dominant condition.

Leiomyomas These neoplasms arise from smooth-muscle components of the intestine and are usually intramural, affecting the overlying mucosa. Ulceration of the mucosa

may cause gastrointestinal hemorrhage of varying severity. Cramping, intermittent abdominal pain is frequently encountered.

Lipomas These tumors occur with greatest frequency in the distal ileum and at the ileocecal valve. They have a characteristic radiolucent appearance, are usually intramural and asymptomatic, but on occasion cause bleeding.

Angiomas While not true neoplasms, these lesions are important because they frequently cause intestinal bleeding. They may take the form of telangiectasia or hemangiomas. Multiple intestinal telangiectasias occur in a nonhereditary form confined to the gastrointestinal tract or as part of the hereditary Osler-Rendu-Weber syndrome. Vascular tumors may also take the form of isolated hemangiomas, most commonly in the jejunum. Angiography, especially during bleeding, is the best procedure for evaluating these lesions.

MALIGNANT TUMORS

While rare, small-bowel malignancies occur in patients with long-standing regional enteritis and celiac sprue as well as in individuals with AIDS. Malignant tumors of the small bowel are frequently associated with fever, weight loss, anorexia, bleeding, and a palpable abdominal mass. After ampullary carcinomas (many of which arise from biliary or pancreatic ducts), the most frequently occurring small-bowel malignancies are adenocarcinomas, lymphomas, carcinoid tumors, and leiomyosarcomas.

Adenocarcinomas The most common primary cancers of the small bowel are adenocarcinomas, accounting for ~50% of malignant tumors. These cancers occur most often in the distal duodenum and proximal jejunum, where they tend to ulcerate and cause hemorrhage or obstruction. Radiologically, they may be confused with chronic duodenal ulcer disease or with Crohn's disease if the patient has long-standing regional enteritis. The diagnosis is best made by endoscopy and biopsy under direct vision. Surgical resection is the treatment of choice.

Lymphomas Lymphoma in the small bowel may be primary or secondary. A diagnosis of a primary intestinal lymphoma requires histologic confirmation in a clinical setting in which palpable adenopathy and hepatosplenomegaly are absent and no evidence of lymphoma is seen on chest radiograph,CT scan, or peripheral blood smear or on bone marrow aspiration and biopsy. Symptoms referable to the small bowel are present, usually accompanied by an anatomically discernible lesion. Secondary lymphoma of the small bowel consists of involvement of the intestine by a lymphoid malignancy extending from involved retroperitoneal or mesenteric lymph nodes (Chap. 112).

Primary intestinal lymphoma accounts for ~20% of malignancies of the small bowel. These neoplasms are non-Hodgkin's lymphomas; they usually have a diffuse, large cell histology and are of T cell origin. Intestinal lymphoma involves the ileum, jejunum, and duodenum, in decreasing frequency, a pattern that mirrors the relative amount of normal lymphoid cells in these anatomic areas. The risk of small-bowel lymphoma is increased in patients with a prior history of malabsorptive conditions (e.g., celiac sprue), regional enteritis, and depressed immune function due to congenital immunodeficiency syndromes, prior organ transplantation, autoimmune disorders, or AIDS.

The development of localized or nodular masses that narrow the lumen results in periumbilical pain (made worse by eating) as well as weight loss, vomiting, and occasional intestinal obstruction. The diagnosis of small-bowel lymphoma may be suspected from the appearance on contrast radiographs of patterns such as infiltration and thickening of mucosal folds, mucosal nodules, areas of irregular ulceration, or stasis of contrast material. The diagnosis can be confirmed by surgical exploration and resection of involved segments. Intestinal lymphoma can occasionally be diagnosed by peroral intestinal mucosal biopsy, but since the disease mainly involves the lamina propria, full-thickness surgical biopsies are usually required.

Resection of the tumor constitutes the initial treatment modality. While postoperative radiation therapy has been given to some patients following a total resection, most authorities favor short-term (three cycles) systemic treatment with combination chemotherapy. The frequent presence of widespread intraabdominal disease at the time of diagnosis and the occasional multicentricity of the tumor often make a total resection impossible. The probability of sustained remission or cure is ~75% in patients with localized disease but is ~25% in individuals with unresectable lymphoma. In patients whose tumors are not resected, chemotherapy may lead to bowel perforation.

A unique form of small-bowel lymphoma, diffusely involving the entire intestine, was first described in oriental Jews and Arabs and is referred to as *immunoproliferative small intestinal disease* (IPSID), *Mediterranean lymphoma,* or*a-heavy chain disease*. This is a B cell tumor. The typical presentation includes chronic diarrhea and steatorrhea associated with vomiting and abdominal cramps; clubbing of the digits may be observed. A curious feature in many patients with IPSID is the presence in the blood and intestinal secretions of an abnormal IgA that contains a shortenedα-heavy chain and is devoid of light chains. It is suspected that the abnormal α chains are produced by plasma cells infiltrating the small bowel. The clinical course of patients with IPSID is generally one of exacerbations and remissions, with death frequently resulting from either progressive malnutrition and wasting or the development of an aggressive lymphoma. The use of oral antibiotics such as tetracycline appears to be beneficial in the early phases of the disorder, suggesting a possible infectious etiology. Combination chemotherapy has been administered during later stages of the disease, with variable results. Results are better when antibiotics and chemotherapy are combined.

Carcinoid Tumors Carcinoid tumors arise from argentaffin cells of the crypts of Lieberkuhn and are found from the distal duodenum to the ascending colon, areas embryologically derived from the midgut. More than 50% of intestinal carcinoids are found in the distal ileum, with most congregating close to the ileocecal valve. Most intestinal carcinoids are asymptomatic and of low malignant potential, but invasion and metastases may occur, leading to the carcinoid syndrome (Chap. 93).

Leiomyosarcomas Leiomyosarcomas often are >5 cm in diameter and may be palpable on abdominal examination. Bleeding, obstruction, and perforation are common.

CANCERS OF THE ANUS

Cancers of the anus account for 1 to 2% of the malignant tumors of the large bowel.

Most such lesions arise in the anal canal, the anatomic area extending from the anorectal ring to a zone approximately halfway between the pectinate (or dentate) line and the anal verge. Carcinomas arising proximal to the pectinate line (i.e., in the transitional zone between the glandular mucosa of the rectum and the squamous epithelium of the distal anus) are known as basaloid, cuboidal, or cloacogenic tumors; about one-third of anal cancers have this histologic pattern. Malignancies arising distal to the pectinate line have a squamous cell histology, ulcerate more frequently, and constitute ~55% of anal cancers. The prognosis for patients with basaloid and squamous cell cancers of the anus is identical when corrected for tumor size and the presence or absence of nodal spread.

The development of anal cancer is associated with infection by human papillomavirus, the same organism etiologically linked to cervical cancer. The virus is sexually transmitted. The infection may lead to anal warts (condyloma accuminata) which may progress to anal intraepithelial neoplasia and on to squamous cell carcinoma. The risk for anal cancer is increased among homosexual males, presumably related to anal intercourse. Anal cancer risk is increased in both men and women with AIDS, possibly because their immunosuppressed state permits more severe papillomavirus infection. Anal cancers occur most commonly in middle-aged persons and are more frequent in women than men. At diagnosis, patients may experience bleeding, pain, sensation of a perianal mass, and pruritus.

Radical surgery (abdominal-perineal resection with lymph node sampling and a permanent colostomy) used to be the treatment of choice for this tumor type. The 5-year survival rate after such a procedure was 55 to 70% in the absence of spread to regional lymph nodes;<20% if nodal involvement was present. An alternative therapeutic approach combining external beam radiation therapy with concomitant chemotherapy has resulted in biopsy-proven disappearance of all tumor in>80% of patients whose initial lesion was <3 cm in size. Tumor has recurred in <10% of these patients, and ~70% of patients with anal cancers can be cured with nonoperative treatment. Surgery should be reserved for the minority of individuals who are found to have residual tumor after being managed initially with radiation therapy combined with chemotherapy.

(Bibliography omitted in Palm version)

Back to Table of Contents

91. TUMORS OF THE LIVER AND BILIARY TRACT - *Jules L. Dienstag*, *Kurt J. Isselbacher*

BENIGN LIVER TUMORS

HEPATOCELLULAR ADENOMAS

Hepatocellular adenomas are benign tumors of the liver found predominantly in women in their third and fourth decades. Their preponderance in women suggests a hormonal influence in their pathogenesis, and oral contraceptives are thought to play an etiologic role. The risk of liver adenomas is increased among those who take anabolic steroids and exogenous androgens. Multiple hepatic adenomas have been associated with glycogen storage disease type I.

Hepatic adenomas occur predominantly in the right lobe of the liver, may be multiple, and are often quite large (>10 cm). Microscopically, they consist of normal or slightly atypical hepatocytes. These cells contain increased glycogen, making them appear paler and larger than normal. Clinical features include pain and the presence of a palpable mass or features of intratumor hemorrhage (pain and circulatory collapse). The diagnosis is usually made by a combination of techniques: sonography, computed tomography (CT), magnetic resonance imaging (MRI), selective hepatic arteriography, and radionuclide scans. The angiographic appearance is typically hypervascular but often also includes hypovascular regions. Technetium 99m scans usually show a defect, because phagocytosing Kupffer cells are absent. Like hepatocellular carcinomas, adenomas have a T_1-intense MRI appearance. The risk of malignant change is small; the risk is higher for large (>10 cm) and multiple adenomas.

Management involves imaging surveillance for small tumors. If the lesion is large (8 to 10 cm), near the surface, and resectable, surgical removal is appropriate. A patient with liver adenoma should stop taking oral contraceptives. Surgical resection may be required for tumors that do not shrink after oral contraceptives are stopped. Pregnancy increases the risk of hemorrhage and should be avoided in women with large adenomas. Patients with multiple large adenomas (e.g., those with glycogen-storage disease) may benefit from liver transplantation.

FOCAL NODULAR HYPERPLASIA

Focal nodular hyperplasia is a benign tumor often identified incidentally on imaging studies or at laparoscopy done for other reasons. Like hepatic adenomas, it occurs predominantly in women; however, oral contraceptives are not implicated, and hemorrhage and necrosis are rare. The risk of hemorrhage, however, appears to be higher in women taking oral contraceptives. Typically, the lesion is a solid tumor, often in the right lobe, with a fibrous core and stellate projections. The fibrous projections contain atypical hepatocytes, biliary epithelium, Kupffer cells, and inflammatory cells. A technetium scan will usually show a hot spot because of the presence of Kupffer cells. The lesion appears vascular on angiography, and septations may be detectable by angiography, helicalCT scan, and, most reliably, byMRI, but only rarely by ultrasound. Surgery is indicated only for symptomatic lesions.

HEMANGIOMA AND OTHER BENIGN TUMORS

Hemangiomas are the most common benign liver tumors, occurring predominantly in women and usually detected incidentally. The prevalence in the general population is in the range of 0.5 to 7.0%. These asymptomatic vascular lesions can be identified byMRI, contrast-enhancedCT, labeled red blood cell nuclide scans, or hepatic angiography. They do not need to be removed unless they are large and are producing a mass effect. Hemorrhage is rare, and malignant change does not occur.

Nodular regenerative hyperplasia consists of multiple hepatic nodules resulting from periportal hepatocyte regeneration with surrounding atrophy. It may be associated with an underlying condition such as malignancy or connective tissue disease. Portal hypertension (in the absence of cirrhosis) is the most common clinical manifestation. Other less common benign hepatic lesions include *bile duct adenomas* and *cystadenomas*.

CARCINOMAS OF THE LIVER

HEPATOCELLULAR CARCINOMA

Epidemiology and Etiology Primary hepatocellular carcinoma is one of the most common tumors in the world. It is especially prevalent in regions of Asia and sub-Saharan Africa, where the annual incidence is up to 500 cases per 100,000 population. In the United States and western Europe, it is much less common; however, the annual incidence in the United States has increased from 1.4/100,000 in the period 1976 to 1980 to 2.4/100,000 in 1991 to 1995. Hepatocellular carcinoma is up to four times more common in men than in women and usually arises in a cirrhotic liver. The incidence peaks in the fifth to sixth decades of life in western countries but one to two decades earlier in regions of Asia and Africa with a high prevalence of liver carcinoma.

The principal reason for the high incidence of hepatocellular carcinoma in parts of Asia and Africa is the frequency of chronic infection with *hepatitis B virus* (HBV) and *hepatitis C virus* (HCV). These chronic infections frequently lead to cirrhosis, which itself is an important risk factor for hepatocellular carcinoma (the risk of liver cancer in a cirrhotic liver is ~3% per year); 60 to 90% of these tumors occur in patients with macronodular cirrhosis. Studies in regions of Asia where hepatocellular carcinoma and HBV infection are prevalent have shown that the incidence of this cancer is about 100-fold higher in individuals with evidence of HBV infection than in noninfected controls. In China, the lifetime risk of developing hepatocellular carcinoma in patients with chronic hepatitis B approaches 40%. In patients with HBV infection and hepatocellular carcinoma, HBV DNA may be integrated into host genomic DNA, both in the tumor cells and in adjacent, uninvolved hepatocytes. In addition, modifications of cellular gene expression occur by insertional mutagenesis, chromosomal rearrangements, or the transcriptional transactivating activity of the X and the pre-S2 regions of the HBV genome.

HCValso leads to hepatocellular carcinoma. HCV genetic material does not become integrated into host genomic DNA. Therefore, the mechanism of HCV carcinogenesis is unclear. In Europe and Japan, HCV appears to be substantially more prevalent than HBVin cases of hepatocellular carcinoma. Both HBV and HCV can be demonstrated in

some patients, but the clinical course of liver malignancy in these patients does not appear to differ from that when only one virus is implicated. One distinction in high-prevalence areas between hepatocellular carcinoma associated with HBV infection and with HCV infection is in the timing of onset. In Asia, HBV is acquired at birth via perinatal transmission, whereas HCV infection is acquired primarily during adulthood from transfused blood and injections. Correspondingly, the onset of liver carcinoma occurs one to two decades earlier in those with lifelong hepatitis B than in persons with adult-acquired hepatitis C. Retrospective analysis indicates that hepatocellular carcinoma occurs on average approximately 30 years after HCV infection and almost exclusively in patients with cirrhosis. The annual incidence of hepatocellular carcinoma in cirrhotic patients with chronic hepatitis C is 1.5 to 4%.

Any agent or factor that contributes to chronic, low-grade liver cell damage and mitosis makes hepatocyte DNA more susceptible to genetic alterations. Thus, as indicated above, *chronic liver disease* of any type is a risk factor and predisposes to the development of liver cell carcinoma. These conditions include alcoholic liver disease,a_1-antitrypsin deficiency, hemochromatosis, and tyrosinemia. In Africa and southern China, *aflatoxin* $B_1$is an important public health hazard. This mycotoxin appears to induce a very specific mutation at codon 249 in the tumor suppressor gene p53.

The loss, inactivation, or mutation of the p53 gene has been implicated in tumorigenesis and is the most common genetic derangement present in human cancers. ThusHBV and aflatoxin B_1 have been implicated in the pathogenesis of hepatocellular carcinoma in regions of Africa and southern China where both agents are prevalent.

In view of the male predominance of liver cancer, hormonal factors may also play a role. Hepatocellular tumors may occur with long-term androgenic steroid administration, with exposure to thorium dioxide or vinyl chloride (see below), and possibly with exposure to estrogens in the form of oral contraceptives.

Clinical and Laboratory Features Cancers of the liver initially may escape clinical recognition because they occur in patients with underlying cirrhosis, and the symptoms and signs may suggest progression of the underlying disease. The most common presenting features are abdominal *pain* with detection of an abdominal mass in the right upper quadrant. There may be a *friction rub* or *bruit* over the liver. Blood-tinged ascites occurs in about 20% of cases. Jaundice is rare, unless there is significant deterioration of liver function or mechanical obstruction of the bile ducts. Serum elevations of alkaline phosphatase and a fetoprotein (AFP) are common (see below). An abnormal type of prothrombin, des-g-carboxy prothrombin, is made and correlates with AFP elevations.

A small percentage of patients with hepatocellular carcinoma have a *paraneoplastic syndrome*; erythrocytosis may result from erythropoietin-like activity produced by the tumor; hypercalcemia may result from secretion of a parathyroid-like hormone. Other manifestations may include hypercholesterolemia, hypoglycemia, acquired porphyria, dysfibrinogenemia, and cryofibrinogenemia.

Imaging procedures to detect liver tumors include ultrasound,CT,MRI, hepatic artery angiography (Chap. 282), and technetium scans. Ultrasound is frequently used to

screen high-risk populations and should be the first test if hepatocellular carcinoma is suspected; it is less costly than scans, is relatively sensitive, and can detect most tumors >3 cm. Helical CT and MRI scans are being used with increasing frequency and have higher sensitivities.

AFPlevels >500 ug/L are found in about 70 to 80% of patients with hepatocellular carcinoma. Lower levels may be found in patients with large metastases from gastric or colonic tumors and in some patients with acute or chronic hepatitis. High levels of serum AFP (>500 to 1000 ug/L) in an adult with liver disease and without an obvious gastrointestinal tumor strongly suggest hepatocellular carcinoma. A rising level suggests progression of the tumor or recurrence after hepatic resection or therapeutic approaches such as chemotherapy or chemoembolization (see below).

Percutaneous *liver biopsy* can be diagnostic if the sample is taken from an area localized by ultrasound orCT. Because these tumors tend to be vascular, percutaneous biopsies should be done with caution. Cytologic examination of ascitic fluid is invariably negative for tumor cells. Occasionally, *laparoscopy* or *minilaparotomy*, to permit liver biopsy under direct vision, may be used. This approach has the additional advantage of sometimes identifying patients who have a localized resectable tumor suitable for partial hepatectomy.

TREATMENT

Staging of hepatocellular carcinoma is based on tumor size (< or > 50% of the liver), ascites (absent or present), bilirubin (<or > 3), and albumin (< or >3) to establish Okuda stages I, II, and III. The Okuda system predicts clinical course better than the American Joint Cancer Commission TNM system. The natural history of each stage without treatment is: stage I, 8 months; stage II, 2 months; stage III, less than 1 month.

The course of *clinically apparent* disease is rapid; if untreated, most patients die within 3 to 6 months of diagnosis. When hepatocellular carcinoma is detected very early by serial screening ofAFP and ultrasound, survival is 1 to 2 years after resection. In selected cases, therapy may prolong life. *Surgical resection* offers the only chance for cure; however, few patients have a resectable tumor at the time of presentation, because of underlying cirrhosis, involvement of both hepatic lobes, or distant metastases (common sites are lung, brain, bone, and adrenal), and the 5-year survival is low. In patients at high risk for the development of hepatocellular carcinoma, screening programs have been initiated to identify small tumors when they are still resectable. Because 20 to 30% of patients with early hepatocellular carcinoma do not have elevated levels of circulating AFP, ultrasonographic screening is recommended as well as AFP determination. In a study in the Far East, persons positive for hepatitis B surface antigen, with or without liver disease, were screened serially; a number of patients with small, subclinical tumors were identified, and surgical resection undertaken. Follow-up observation revealed a 5-year survival rate in this group of 70% and a 10-year survival rate of 50%. These Asian patients, however, were unusual in that they had minimal or no liver disease and their tumors tended to be unifocal or encapsulated. The findings are in contrast to a study in a large population of Italian patients with cirrhosis, associated in most cases with chronicHBVand/orHCVinfections; screening every 3 to 12 months permitted the detection of a 3% annual incidence of

cancer in this cohort but in most cases failed to achieve the goal of early detection of surgically treatable disease. No randomized study has yet shown survival benefit for screening patients at high risk of developing hepatocellular carcinoma.

Liver transplantation may be considered as a therapeutic option; tumor recurrence or metastases are the major problems. Patients who have a single lesion£5 cm or three or fewer lesions£3 cm have survival after liver transplantation that is the same as survival after transplantation for nonmalignant liver disease (Chap. 301). Other approaches include (1) hepatic artery embolization and chemotherapy (chemoembolization), (2) alcohol or radio-frequency ablation via ultrasound-guided percutaneous injection, and (3) ultrasound-guided cryoablation.

Treatment options for unresectable disease are limited. Randomized trials have not shown a survival advantage after chemoembolization. The liver cannot tolerate high doses of radiation. The disease is not responsive to chemotherapy, including newer agents such as gemcytabine. Investigative immunotherapy and gene therapy techniques have not yet been successful. Based on the presence of hormone receptors on the tumor, tamoxifen has been tested, but without success, and octreotide has had some modest activity. In patients with resectable tumors, polyprenoic acid (a retinoic acid formulation) and intraarterial ^{131}I-labeled lipiodol have been reported to reduce the rate of recurrence.

Prevention is the preferred strategy. Hepatitis B vaccine can prevent infection and its sequelae, and a reduction in hepatocellular carcinoma has been seen in Taiwan with the introduction of universal vaccination of children. Interferon treatment reduces the incidence of hepatic failure, death, and liver cancer in patients infected withHBV. Treatment with interferon may lower the risk of development of liver cancer in patients with hepatitis C-related cirrhosis (Chap. 297), but additional studies are needed.

OTHER MALIGNANT LIVER TUMORS

Fibrolamellar carcinoma differs from the typical hepatocellular carcinoma in that it tends to occur in young adults without underlying cirrhosis. This tumor is nonencapsulated but well circumscribed and contains fibrous lamellae; it grows slowly and is associated with a longer survival if treated. Surgical resection has resulted in 5-year survivals >50%; if the lesion is nonresectable, liver transplantation is an option, and the outcome far exceeds that observed in the nonfibrolamellar variety of liver cancer. *Hepatoblastoma* is a tumor of infancy that typically is associated with very high serumAFPlevels. The lesions are usually solitary, may be resectable, and have a better 5-year survival than that of hepatocellular carcinoma. *Angiosarcoma* consists of vascular spaces lined by malignant endothelial cells. Etiologic factors include prior exposure to thorium dioxide (Thorotrast), polyvinyl chloride, arsenic, and androgenic anabolic steroids. *Epithelioid hemangioendothelioma* is of borderline malignancy; most cases are benign, but bone and lung metastases occur. This tumor occurs in early adulthood, presents with right upper quadrant pain, is heterogeneous on sonography, hypodense onCT, and without neovascularity on angiography. Immunohistochemical staining reveals expression of factor VIII antigen. In the absence of extrahepatic metastases, these lesions can be treated by surgical resection or liver transplantation.

METASTATIC TUMORS

Metastatic tumors of the liver are common, ranking second only to cirrhosis as a cause of fatal liver disease. In the United States, the incidence of metastatic carcinoma is at least 20 times greater than that of primary carcinoma. At autopsy, hepatic metastases occur in 30 to 50% of patients dying from malignant disease.

Pathogenesis The liver is uniquely vulnerable to invasion by tumor cells. Its size, high rate of blood flow, double perfusion by the hepatic artery and portal vein, and its Kupffer cell filtration function combine to make it the next most common site of metastases after the lymph nodes. In addition, local tissue factors or endothelial membrane characteristics appear to enhance metastatic implants. Virtually all types of neoplasms except those primary in the brain may metastasize to the liver. The most common primary tumors are those of the gastrointestinal tract, lung, and breast, as well as melanomas. Less common are metastases from tumors of the thyroid, prostate, and skin.

Clinical Features Most patients with metastases to the liver present with symptoms referable only to the primary tumor, and the asymptomatic hepatic involvement is discovered in the course of clinical evaluation. Sometimes hepatic involvement is reflected by nonspecific symptoms of weakness, weight loss, fever, sweating, and loss of appetite. Rarely, features indicating active hepatic disease, especially abdominal pain, hepatomegaly, or ascites, are present. Patients with widespread metastatic liver involvement usually have suggestive clinical signs of cancer and hepatic enlargement. Some have localized induration or tenderness, and, occasionally, a friction rub may be found over tender areas of the liver.

Results of liver biochemical tests are often abnormal, but the elevations in marker levels are often only mild and nonspecific. These signs reflect the effects of fever and wasting as well as those of the infiltrating neoplastic process itself. An increase in serum alkaline phosphatase is the most common and frequently the only abnormality. Hypoalbuminemia, anemia, and occasionally a mild elevation of aminotransferase levels may also be found with more widespread disease. Substantially elevated serum levels of carcinoembryonic antigen are usually found when the metastases are from primary malignancies in the gastrointestinal tract, breast, or lung.

Diagnosis Evidence of metastatic invasion of the liver should be sought actively in any patient with a primary malignancy, especially of the lung, gastrointestinal tract, or breast, before resection of the primary lesion. An elevated level of alkaline phosphatase or a mass apparent on ultrasound,CT, orMRIexamination of the liver may provide a presumptive diagnosis. Blind percutaneous needle biopsy of the liver will result in a positive diagnosis of metastatic disease in only 60 to 80% of cases with hepatomegaly and elevated alkaline phosphatase levels. Serial sectioning of specimens, two or three repeated biopsies, or cytologic examination of biopsy smears may increase the diagnostic yield by 10 to 15%. The yield is increased when biopsies are directed by ultrasound or CT or obtained during laparoscopy.

TREATMENT

Most metastatic carcinomas respond poorly to all forms of treatment, which is usually only palliative. Rarely a single, large metastasis can be removed surgically. Systemic chemotherapy may slow tumor growth and reduce symptoms, but it does not alter the prognosis. Chemoembolization, intrahepatic chemotherapy, and alcohol or radio-frequency ablation may provide palliation.

CHOLANGIOCARCINOMA

Benign tumors of the extrahepatic bile ducts are extremely rare causes of mechanical biliary obstruction. Most of these are papillomas, adenomas, or cystadenomas and present with obstructive jaundice or hemobilia. Adenocarcinoma of the extrahepatic ducts is more common. There is a slight male preponderance (60%), and the incidence peaks in the fifth to seventh decades. Apparent predisposing factors include (1) some chronic hepatobiliary parasitic infestations, (2) congenital anomalies with ectactic ducts, (3) sclerosing cholangitis and chronic ulcerative colitis, and (4) occupational exposure to possible biliary tract carcinogens (employment in rubber or automotive plants). Cholelithiasis is not clearly a predisposing factor for cholangiocarcinoma. The lesions of cholangiocarcinoma may be diffuse or nodular. Nodular lesions often arise at the bifurcation of the common bile duct (Klatskin tumors) and are usually associated with a *collapsed gallbladder*, a finding that mandates cholangiography to view proximal hepatic ducts.

Patients with cholangiocarcinoma usually present with biliary obstruction, painless jaundice, pruritus, weight loss, and acholic stools. A deep-seated, vaguely localized right upper quadrant pain may be noted. Hepatomegaly and a palpable, distended gallbladder (unless the lesion is high in the duct) are frequent accompanying signs. Fever is unusual unless associated with ascending cholangitis. Because the obstructing process is gradual, the cholangiocarcinoma is often far advanced by the time it presents clinically. The diagnosis is most frequently made by cholangiography following ultrasound demonstration of dilated intrahepatic bile ducts. Any focal strictures of the bile ducts should be considered malignant until proved otherwise. Endoscopic cholangiography permits obtaining specimens for cytology (sensitivity ~60%) and insertion of stents for biliary drainage. Survival of 1 to 2 years is possible in some cases. Perhaps 20% of patients have surgically resectable tumors, but 5-year survival is only 10 to 30%. The high recurrence rate limits the value of liver transplantation. Photodynamic therapy (intravenous hematoporphyrin with cholangioscopically delivered light) has been used with promising early results.

CARCINOMA OF THE PAPILLA OF VATER

The ampulla of Vater may be involved by extension of tumor arising elsewhere in the duodenum or may itself be the site of origin of a sarcoma, carcinoid tumor, or adenocarcinoma. Papillary adenocarcinomas are associated with slow growth and a more favorable clinical prognosis than diffuse, infiltrative cancers of the ampulla, which are more frequently widely invasive. The presenting clinical manifestation is usually obstructive jaundice. Endoscopic retrograde cannulation of the pancreatic duct is the preferred diagnostic technique when ampullary carcinoma is suspected, because it allows for direct endoscopic inspection and biopsy of the ampulla and for pancreatography to exclude a pancreatic malignancy. Cancer of the papilla is usually

treated by wide surgical excision. Lymph node or other metastases are present at the time of surgery in approximately 20% of cases, and the 5-year survival rate following surgical therapy in this group is only 5 to 10%. In the absence of metastases, radical pancreaticoduodenectomy (the Whipple procedure) is associated with 5-year survival rates as high as 40%.

CANCER OF THE GALLBLADDER

Most cancers of the gallbladder develop in conjunction with stones rather than polyps. In patients with gallstones, the risk for developing gallbladder cancer, while increased, is still quite low. In one study, gallbladder cancer developed in only 5 of 2583 patients with gallstones followed for a median of 13 years. In the United States, adenocarcinomas make up the vast majority of the estimated 6500 new cases of gallbladder cancer diagnosed each year. The female/male ratio is 4:1, and the mean age at diagnosis is approximately 70 years. The clinical presentation is most often one of unremitting right upper quadrant pain associated with weight loss, jaundice, and a palpable right upper quadrant mass. Cholangitis may supervene. The preoperative diagnosis of the condition has been facilitated by ultrasound andCT. CT is also useful in guiding fine-needle aspiration and biopsy.

Once symptoms have appeared, spread of the tumor outside the gallbladder by direct extension or by lymphatic or hematogenous routes is almost invariable. Over 75% of gallbladder carcinomas are unresectable at the time of surgery, the exceptions being tumors discovered incidentally at laparotomy. If the tumor is found by the pathologist, no additional therapy is required. If the tumor is noted by the surgeon on routine cholecystectomy, a second operation is generally performed to resect the adjacent liver, bile duct, and local lymph nodes. Incidental resectable gallbladder tumors have a 50% 5-year survival. The 1-year mortality rate for unresectable disease is about 95%, and <5% of patients survive 5 years. Radical operative resection does not appear to improve survival. Trials of radiation and chemotherapy in patients with gallbladder cancer have been disappointing.

(Bibliography omitted in Palm version)

Back to Table of Contents

92. PANCREATIC CANCER - *Robert J. Mayer*

INCIDENCE AND ETIOLOGY

The incidence of pancreatic carcinoma in the United States has increased significantly as the median life expectancy of the American population has lengthened. The tumor results in the death of >98% of afflicted patients. 28,200 individuals died of pancreatic cancer in 2000, making it the fifth most common cause of cancer-related mortality. The disease is more common in males than in females and in blacks than in whites. It rarely develops before the age of 50.

Little is known about the causes of pancreatic cancer. Cigarette smoking is the most consistent risk factor, with the disease being two to three times more common in heavy smokers than in nonsmokers. Whether this association is due to a direct carcinogenic effect of tobacco metabolites on the pancreas or an as yet undefined exposure that occurs more frequently in cigarette smokers is uncertain. Patients with chronic pancreatitis are at increased risk of pancreatic cancer, as are persons with long-standing diabetes mellitus. Obesity is a risk factor for pancreatic cancer; risk is directly related to increased calorie intake. Alcohol abuse or cholelithiasis are not risk factors for pancreatic cancer. Nor is pancreatic cancer associated with coffee consumption. Mutations in K-*ras* genes have been found in >85% of specimens of human pancreatic cancer. Pancreatic cancer has been associated with mutation of the *p16*$_{INK4}$gene located on chromosome 9p21, a gene also implicated in the pathogenesis of malignant melanoma.

CLINICAL FEATURES

More than 90% of pancreatic cancers are ductal adenocarcinomas, with islet cell tumors constituting the remaining 5 to 10%. Pancreatic cancers occur twice as frequently in the pancreatic head (70% of cases) as in the body (20%) or tail (10%) of the gland.

With the exception of jaundice, the initial symptoms associated with pancreatic cancer are often insidious and are usually present for >2 months before the cancer is diagnosed (Table 92-1). Pain and weight loss are present in>75% of patients. The pain typically has a gnawing, visceral quality, occasionally radiating from the epigastrium to the back. Pain is often a more severe problem in lesions arising in the body or tail of the gland, as such tumors may become quite large before being detected. Characteristically, the pain improves somewhat when the patient bends forward. The development of significant pain suggests retroperitoneal invasion and infiltration of the splanchnic nerves, indicating that the primary lesion is advanced and is not surgically resectable. Rarely, such pain may be transient and associated with hyperamylasemia, indicative of acute pancreatitis caused by ductal obstruction by tumor. The weight loss observed in most patients is primarily the result of anorexia, although in the initial period of the disease, subclinical malabsorption may also be a contributing factor.

Jaundice due to biliary obstruction is found in >80% of patients having tumors in the pancreatic head and is typically accompanied by dark urine, a claylike appearance of stool, and pruritus. In contrast to the "painless jaundice" sometimes observed in patients having carcinomas of the bile ducts, duodenum, or periampullary regions, most icteric

individuals with ductal carcinomas of the pancreatic head will complain of significant abdominal discomfort. Although the gallbladder is usually enlarged in patients with carcinoma of the head of the pancreas, it is palpable in <50% (Courvoisier's sign). However, the presence of an enlarged gallbladder in a jaundiced patient without biliary colic should suggest malignant obstruction of the extrahepatic biliary tree.

Glucose intolerance, presumably a direct consequence of the tumor, often develops within 2 years of the clinical diagnosis. Other initial manifestations include venous thrombosis and migratory thrombophlebitis (Trousseau's syndrome), gastrointestinal hemorrhage from varices due to compression of the portal venous system by tumor, and splenomegaly caused by cancerous encasement of the splenic vein.

DIAGNOSTIC PROCEDURES (Fig. 92-1)

Despite the availability of serologic tests for tumor-associated antigens, such as the carcinoembryonic antigen (CEA) and CA 19-9, and noninvasive imaging techniques, such as computed tomography (CT) and ultrasonography, the early diagnosis of a potentially resectable pancreatic carcinoma remains extremely difficult. The nonspecificity of the initial symptoms and the poor sensitivity of both serologic assays and noninvasive techniques have frustrated the development of effective screening procedures. When the disease is clinically suspected in a patient having vague, persistent abdominal complaints, ultrasound should be performed to visualize the gallbladder and the pancreas, as well as upper gastrointestinal contrast radiographs to rule out a hiatal hernia or a peptic ulcer. If these studies fail to provide an explanation for the symptoms, a CT scan should be considered. It should encompass not only the pancreas but also the liver, retroperitoneal lymph nodes, and pelvis, as pancreatic cancer frequently spreads within the abdomen. While more costly than ultrasonography, CT is technically simpler, more reproducible, provides better definition of the body and tail of the pancreas, and requires less interpretive skill. CT generally detects a malignant pancreatic lesion in>80% of cases; in 5 to 15% of patients with proven pancreatic carcinoma, the CT scan shows only generalized pancreatic enlargement suggesting pancreatitis rather than malignancy. False-positive results occur in about 5 to 10% of cases where no tumor was found on laparotomy. Magnetic resonance imaging (MRI) has not been shown to be better than CT in the evaluation of pancreatic lesions. The value of positron emission tomography (PET) has not been defined. When clinical circumstances dictate additional diagnostic evaluation, endoscopic retrograde cholangiopancreatography (ERCP) with endoscopic ultrasonography (EUS) may clarify the cause of ambiguous CT or ultrasound findings. The characteristic findings are stenosis or obstruction of either the pancreatic or the common bile duct; both duct systems are abnormal in over half the cases. Carcinoma and chronic pancreatitis can be difficult to distinguish by ERCP, particularly if both diseases are present. False-negative results with ERCP are infrequent (<5%) and usually occur in the setting of islet cell, rather than ductal, carcinomas.

Selective and superselective angiography may be of value in some patients. Angiography is an effective means of detecting carcinomas in the body and tail of the pancreas by the demonstration of vascular narrowing, displacement, or occlusion by tumor. Angiography is being replaced as a diagnostic and staging procedure by spiralCTscanning with contrast imaging. This high-resolution technology predicts the

resectability of the tumor if no disease is found outside the pancreas, obstruction of the superior mesenteric-portal vein confluence is absent, or tumor extension to the celiac axis and superior mesenteric arteries is not found. Radiographic staging criteria are shown inTable 92-2.

Regardless of the results of the above diagnostic studies, a histologic confirmation of pancreatic cancer is mandatory; similar findings can result from other neoplasms such as islet cell tumor or lymphoma, for which the therapeutic approach and prognosis differ from those for ductal carcinoma. In patients with unresectable disease or medical contraindications to surgical resection, tissue may be obtained through a percutaneous needle aspiration biopsy of the pancreas withCT or ultrasonographic guidance.

Unfortunately, however, even laparotomy may not provide a definitive diagnosis, because chronic pancreatitis may also produce a hard mass in the head of the pancreas indistinguishable from carcinoma by palpation. Furthermore, a superficial biopsy of such a mass may not show neoplastic tissue, revealing only evidence of pancreatitis, as the cancer is often surrounded by edematous, inflamed, and fibrotic tissue (i.e., chronic pancreatitis).

TREATMENT

Complete surgical resection of pancreatic tumors offers the only effective treatment for this disease. Unfortunately, such "curative" operations are only possible in 10 to 15% of patients with pancreatic cancer, usually those individuals with a tumor in the pancreatic head in whom jaundice was the initial symptom. Patients considered for such a procedure should have no evidence of metastatic spread on a chest radiograph and abdominal-pelvicCT scan and should be operated on by an experienced surgeon, as mortality rates of>15% have been associated with this procedure. Curative resection is usually preceded by laparoscopic inspection of the abdomen to confirm absence of occult disease spread to the omentum, peritoneum, or liver, which would preclude curative resection. Although the potential for cure in patients with pancreatic cancer is restricted to the few who are able to undergo a complete surgical resection, the 5-year survival rate following such operations is only 10%. Nonetheless, the procedure is worth attempting, particularly for lesions in the pancreatic head, since ductal carcinomas often cannot be distinguished preoperatively from ampullary, duodenal, and distal bile duct tumors or pancreatic cyst adenocarcinomas, all of which have far higher rates of resectability and cure. Furthermore, patients who undergo resection and eventually experience disease recurrence survive three to four times longer than those whose tumor is not excised, indicating that such operations have a palliative effect. The risk for tumor recurrence is not affected by the type of operative procedure -- i.e., total pancreatectomy versus pancreaticoduodenectomy ("Whipple resection") -- but it is increased by the presence of lymph node metastases or tumor invasion into adjacent viscera. As a rule, pancreaticoduodenectomy or distal pancreatectomy seems preferable to total pancreatectomy because of the retention of exocrine function and avoidance of brittle diabetes.

The median survival for patients whose pancreatic cancers are surgically unresectable is 6 months. Management is directed at palliation of symptoms. Ambulatory patients having tumors in the pancreatic head should be considered for surgical diversion of the

biliary system. If jaundice has already developed, therapeutic options include either nonoperative biliary decompression by endoscopic or percutaneous, transhepatic biliary drainage or surgical biliary bypass. External beam radiation in patients with unresectable tumors that have not spread beyond the pancreas does not appear to prolong survival, although a sufficient reduction in tumor size may lead to palliation of pain. However, the addition of chemotherapy with fluorouracil (5-FU) to external beam irradiation has increased the survival time for these patients, perhaps because 5-FU acts as a radiosensitizing agent. In a small patient population, a similar combination of radiation therapy and 5-FU appears to have prolonged the survival and increased the cure rate as compared to a prospectively randomized nontreatment control group of patients who had a complete surgical resection of their pancreatic cancer. This observation needs to be confirmed before it can be accepted. The possibility of administering such chemoradiation therapy at diagnosis and before surgery ("neoadjuvant" treatment), to increase the potential for resectability, is under investigation. Intraoperative radiation therapy has the potential to deliver higher doses of radiation to the tumor while sparing neighboring tissues but does not give better results than external beam treatment.

Chemotherapy in the management of patients with widely metastatic pancreatic cancer has been disappointing. Gemcitabine, a deoxycytidine analogue, produces improvement in the quality of life for patients with advanced pancreatic cancer. However, duration of survival is only moderately improved. Newer forms of treatment, including combining gemcitabine with other cytotoxic agents or therapies directed at specific molecular targets, such as K-*ras*, or p53 are being evaluated. Experimental therapy should constitute the initial treatment for consenting, ambulatory patients.**Pancreatic endocrine tumors are discussed in Chap. 93.*

(Bibliography omitted in Palm version)

Back to Table of Contents

93. ENDOCRINE TUMORS OF THE GASTROINTESTINAL TRACT AND PANCREAS - *Robert T. Jensen*

Gastrointestinal neuroendocrine tumors (NETs) are derived from the diffuse neuroendocrine system of the gastrointestinal (GI) tract, which is composed of amine- and acid-producing cells with different hormonal profiles, depending on the site of origin. The tumors they produce can be divided into carcinoid tumors and pancreatic endocrine tumors (PETs). These tumors were originally classified as APUDomas (for *a*mine *p*recursor *u*ptake and *d*ecarboxylation), as were pheochromocytomas, melanomas, and medullary thyroid carcinomas because they share certain cytochemical, pathologic, and biologic features (Table 93-1). APUDomas were thought to have a similar embryonic origin from neural crest cells, but the peptide-secreting cells are not of neuroectodermal origin.

CLASSIFICATION, PATHOLOGY, AND TUMOR BIOLOGY OF NETS

NETs are generally composed of monotonous sheets of small round cells with uniform nuclei; mitoses are uncommon. They can be tentatively identified on routine histology; however, these tumors are principally recognized by their histologic staining patterns due to shared cellular proteins. Historically, silver staining was used, and tumors were classified as showing an argentaffin reaction if they took up and reduced silver or as being argyrophilic if they did not reduce it. Immunocytochemical localization of chromogranins (A,B,C), neuron-specific enolase, or synaptophysin, which are all neuroendocrine cell markers, are now used (Table 93-1). Chromogranin A is the most widely used.

Ultrastructurally, these tumors possess electron-dense neurosecretory granules and frequently contain small clear vesicles that correspond to synaptic vesicles of neurons.NETssynthesize numerous peptides, growth factors, and bioactive amines that may be ectopically secreted, giving rise to a specific clinical syndrome (Table 93-2). The diagnosis of the specific syndrome, such as a VIPoma [vasoactive intestinal peptide (VIP)-secreting tumor], requires the clinical features of the disease and cannot be made from the immunocytochemistry results only (Table 93-1). Furthermore, pathologists cannot distinguish between benign and malignant NETs unless metastases or invasion are present.

Carcinoid tumors are frequently classified according to their anatomic area of origin (i.e., foregut, midgut, hindgut) because tumors with similar areas of origin share functional manifestations, histochemistry, and secretory products (Table 93-3). Foregut tumors generally have a low serotonin [5-hydroxytryptamine (5-HT)] content, are argentaffin-negative but argyrophilic, occasionally secrete adrenocorticotropic hormone (ACTH) or 5-hydroxytryptophan (5-HTP) causing an atypical carcinoid syndrome (Fig. 93-1), make several hormones, and may metastasize to bone. They uncommonly produce a clinical syndrome due to the secreted products. Midgut carcinoids are argentaffin-positive, have a high serotonin content, most frequently cause the typical carcinoid syndrome when they metastasize (Table 93-3,Fig. 93-1), release serotonin and tachykinins (substance P, neuropeptide K, substance K), rarely secrete 5-HTP or ACTH, and uncommonly metastasize to bone. Hindgut carcinoids (rectum and transverse and descending colon) are argentaffin-negative, often argyrophilic, rarely

contain serotonin or cause the carcinoid syndrome, rarely secrete 5-HTP or ACTH, contain numerous peptides, and may metastasize to bone.

PETscan be classified into specific functional or nonfunctional syndromes (Table 93-2). Each of the functional syndromes is associated with symptoms due to the specific hormone released. In contrast, nonfunctional PETs release no products that cause a specific clinical syndrome. "Nonfunctional" is a misnomer in the strict sense because these tumors frequently ectopically secrete a number of peptides [pancreatic polypeptide (PP) chromogranin A, neurotensin]; however, they cause no specific clinical syndrome. The symptoms caused by nonfunctional PETs are entirely due to the tumor per se.

Carcinoid tumors can occur in almost any GI tissue (Table 93-3); however, most (70%) originate from one of four sites; bronchus, jejunum/ileum, rectum, or appendix. In the past, carcinoid tumors most frequently occurred in the appendix (i.e., 40%); however, the bronchus/lung is now the most common site (32%). Overall, the GI tract is the most common site for these tumors, comprising 74%, with the respiratory tract a distant second at 25%.

The term *pancreatic endocrine tumor*, although widely used, is also a misnomer, because these tumors can occur either almost exclusively in the pancreas (insulinomas, glucagonomas, nonfunctionalPETs, PETs causing hypercalcemia) or at both pancreatic and extrapancreatic sites [gastrinomas,VIPomas, somatostatinomas, GRFomas (GRF, growth hormone-releasing factor)]. PETs are also called *islet cell tumors*; however, this term is discouraged because it is not established that many originate from the islets and many can occur at extrapancreatic sites.

The exact incidence of carcinoid tumors orPETsvaries according to whether only symptomatic or all tumors are considered. The incidence of clinically significant carcinoids is 7 to 13 cases per million population per year, whereas malignant carcinoids are reported at autopsy in 21 to 84 cases per million population per year. Clinically significant PETs have a prevalence of 10 cases per million population, with insulinomas, gastrinomas, and nonfunctional PETs having an incidence of 0.5 to 2 cases per million population per year (Table 93-2); VIPomas are 2- to 8-fold less common, glucagonomas are 17- to 30-fold less common, and somatostatinomas the least common. In autopsy studies 0.5 to 1.5% of all cases have a PET; however, in fewer than 1 in 1000 cases was a functional tumor present.

Both carcinoid tumors andPETscommonly show malignant behavior (Tables 93-2,93-3). With PETs, except for insulinomas in which <10% are malignant, 50 to 100% are malignant. With carcinoid tumors the percentage showing malignant behavior varies in different locations. For the four most common sites of occurrence the incidence of metastases varies greatly: jejunum/ileum (70%) > appendix (35%)> lung/bronchus (27%)> rectum (14%). A number of factors are important in determining survival and the aggressiveness of the tumor (Table 93-4). The presence of liver metastases is the single most important prognostic factor for both carcinoid tumors and PETs. The size of the primary tumor is particularly important in the development of liver metastases. For example, with small-intestinal carcinoids, which are the most frequent cause of the carcinoid syndrome due to metastatic disease in the liver (Table 93-2), metastases

occur in 15 to 25% if the tumor diameter is <1 cm, in 58 to 80% if it is 1 to 2 cm and in >75% if it is>2 cm. The size of the primary tumor has also been shown to be an independent predictor of the development of liver metastases for gastrinomas and other PETs. The presence of lymph node metastases, the depth of invasion, various histologic features (differentiation, mitotic rates, growth indices), and flow cytometric results such as the presence of aneuploidy are all important prognostic factors for the development of metastatic disease. The development of the carcinoid syndrome, older age, male gender, the presence of a symptomatic tumor, or greater increases in a number of tumor markers [5-hydroxyindolacetic acid (5-HIAA), neuropeptide K, chromogranin A] also adversely affect prognosis in patients with carcinoid tumors (Table 93-4). With PETs or gastrinomas, a worse prognosis is associated with female gender, overexpression of the *Ha-Ras* oncogene or p53, the absence of multiple endocrine neoplasia (MEN) type 1, and higher levels of various tumor markers (e.g., chromogranin A, gastrin).

A number of genetic disorders are associated with an increased incidence of neuroendocrine tumors (Table 93-5). Each one is caused by a loss of a possible tumor suppressor gene. The most important isMEN-1, an autosomal dominant disorder due to a defect in a 10-exon gene on 11q13, which encodes for a 610-amino-acid nuclear protein, menin (Chap. 339). In patients with MEN-1, 95 to 100% develop hyperparathyroidism due to parathyroid hyperplasia, 80 to 100% develop nonfunctionalPETs, 54 to 80% develop pituitary adenomas, and bronchial carcinoids develop in 8%, thymic carcinoids in 8%, and gastric carcinoids in 13 to 30% of patients with Zollinger-Ellison syndrome. Functional PETs occur in 80% of patients with MEN-1, with 54% developing Zollinger-Ellison syndrome, 21% insulinomas, 3% glucagonomas, and 1%VIPomas. MEN-1 is present in 20 to 25% of all patients with Zollinger-Ellison syndrome, in 4% of those with insulinomas, and in <5% of those with other PETs.

Three phakomatoses associated withNETs are von Hippel-Lindau disease, von Recklinghausen's disease, or neurofibromatosis type 1 (NF-1), and tuberous sclerosis (Bourneville's disease). Von Hippel-Lindau disease is an autosomal dominant disorder due to defects in a gene on chromosome 3p25, which encodes a 213-amino-acid protein that interacts with the elongin family of proteins as a transcriptional regulator and in abnormal protein destruction (Chap. 370). In addition to cerebellar hemangioblastomas, renal cancer, and pheochromocytomas, 10 to 17% of these patients develop aPET. Most are nonfunctional, although insulinomas andVIPomas are reported. Patients with NF-1 have defects in a gene on chromosome 17q11.2 encoding for a 2845-amino-acid protein, neurofibromin, which functions in normal cells as a suppressor of the Ras signaling cascade (Chap. 370). Up to 12% of these patients develop an upper GI carcinoid tumor, characteristically in the periampullary region (54%). Many of such tumors are classified as somatostatinomas because they contain somatostatin immunocytochemically; however, they seldom produce a clinical somatostatinoma syndrome. NF-1 has rarely been associated with insulinomas and Zollinger-Ellison syndrome. Tuberous sclerosis is caused by mutations that alter either the 1164-amino-acid protein, hamartin (TSC1), or the 1807-amino-acid protein, tuberin (TSC2) (Chap. 370). Both hamartin and tuberin interact in a pathway related to cytosolic G protein regulation. A few cases including nonfunctional and functional PETs (insulinomas and gastrinomas) have been reported (Table 93-5).

Changes in theMEN-1 gene, p16/MTS1 tumor suppressor gene, and DPC 4/Smad 4 gene; amplification of the HER-2/neu protooncogene; and deletions of unknown tumor suppressor genes on chromosomes 1 and 3p may be important in the pathogenesis ofPETs. Loss of heterozygosity at the MEN-1 locus on chromosome 11q13 has been found in 93% of sporadic PETs (patients without MEN-1) and in 26 to 75% of sporadic carcinoid tumors. Mutations in the MEN-1 gene were found in 31 to 34% of sporadic gastrinomas.

CARCINOID TUMORS AND CARCINOID SYNDROME

GENERAL TUMOR CHARACTERISTICS OF THE MOST COMMON GI CARCINOID TUMORS

Appendiceal Carcinoids These occur in 1 in every 200 to 300 appendectomies, usually in the appendiceal tip. More than 90% are <1 cm in diameter; 35% have metastases. In one study of 1570 appendiceal carcinoids, 62% were localized and 27% had regional and 8% distant metastases. Half of the carcinoids between 1 and 2 cm in diameter had metastasized to lymph nodes.

Small-Intestinal Carcinoids These are frequently multiple. Between 70 and 80% are present in the ileum and 70% within 60 cm (24 in.) of the ileocecal valve; 40% are<1 cm in diameter, 32% are 1 to 2 cm, and 29% are>2 cm; and 35 to 70% are associated with metastases. They characteristically cause a marked fibrotic reaction, which can lead to intestinal obstruction. Distant metastases occur to liver in 36 to 60% of patients, to bone in 3%, and to lung in 4%. Between 15 and 25% of small (<1 cm) carcinoid tumors of the small intestine have metastases; 58 to 100% of tumors 1 to 2 cm in diameter have metastases. Carcinoids also occur in the duodenum, with 21% having metastases. In one study, no duodenal tumor <1 cm metastasized, whereas 33% of those >2 cm had metastases. Small-intestinal carcinoids are the most common cause (60 to 87%) of the carcinoid syndrome (Table 93-6).

Rectal Carcinoids Rectal carcinoids are found in 1 of every 2500 proctoscopies. Nearly all occur 4 to 13 cm above the dentate line. Most are small, with 66 to 80% being <1 cm in diameter; 5% metastasize. Tumors 1 to 2 cm in diameter metastasize in 5 to 30% of patients, and tumors>2 cm, which are uncommon, in>70%.

Bronchial Carcinoids Bronchial carcinoids are not related to smoking. A number of different classifications of bronchial carcinoid tumors are proposed. In some studies lungNETs are classified into four categories: typical carcinoid (also called bronchial carcinoid tumor, Kulchitsky cell carcinoma-I, KCC-I); atypical carcinoid (also called well-differentiated neuroendocrine carcinoma, KC-II); intermediate small cell neuroendocrine carcinoma; and small cell neuroendocarcinoma (KC-III). Another proposed classification includes three categories of lung NETs: benign or low-grade malignant (typical carcinoid); low-grade malignant (atypical carcinoid), and high-grade malignant (poorly differentiated carcinoma of the large cell or small cell type). These different categories of lung NETs have different prognoses, varying from excellent for typical carcinoid to poor for small cell neuroendocrine carcinomas.

Gastric Carcinoids These account for 3 of every 1000 gastric neoplasms. Three

different subtypes of gastric carcinoids are noted. Each originates from gastric enterochromaffin-like (ECL) cells in the gastric mucosa. Two subtypes are associated with hypergastrinemic states: (1) chronic atrophic gastritis (type I) (80% of all gastric carcinoids); and (2) Zollinger-Ellison syndrome, almost always as part of theMEN-1 syndrome (type II) (6% of all cases). These tumors generally pursue a benign course, with 9 to 30% associated with metastases. They are usually multiple and small and infiltrate only to the submucosa. The third subtype of gastric carcinoid (type III) (sporadic) occurs without hypergastrinemia (14% of all carcinoids) and pursues an aggressive course, with 54 to 66% developing metastases. Sporadic carcinoids are usually single, large tumors; 50% have atypical histology and can be a cause of the carcinoid syndrome.

CARCINOID TUMORS WITHOUT THE CARCINOID SYNDROME

The age of patients at diagnosis ranges from 10 to 93 years, with a mean age of 63 years for carcinoid tumors of the small intestine and 66 years for those of the rectum. The presentation is diverse and related to the site of origin and extent of malignant spread. In the appendix, carcinoid tumors are usually found incidentally during surgery for suspected appendicitis. Small-intestinal carcinoids in the jejunum/ileum present with periodic abdominal pain (51%), intestinal obstruction with ileus/invagination (31%), an abdominal tumor (17%), or GI bleeding (11%). Because of the vagueness of the symptoms the diagnosis is usually delayed approximately 2 years from onset of the symptoms, with a range up to 20 years. Duodenal, gastric, and rectal carcinoids are most frequently found by chance at endoscopy. The most common symptoms of rectal carcinoids are melena/bleeding (39%), constipation (17%), and diarrhea (12%). Bronchial carcinoids are frequently discovered as a lesion on a chest radiograph, and 31% of the patients are asymptomatic. Thymic carcinoids present as anterior mediastinal masses, usually on chest radiograph of computed tomography (CT) scan. Ovarian and testicular carcinoids usually present as masses discovered on physical examination or by ultrasound. Metastatic carcinoid tumor in the liver presents frequently as hepatomegaly in a patient who may have minimal symptoms and near-normal liver function test results.

CARCINOID TUMORS WITH SYSTEMIC SYMPTOMS DUE TO SECRETED PRODUCTS

Carcinoid tumors can contain numerous GI peptides: gastrin, insulin, somatostatin, motilin, neurotensin, tachykinins (substance K, substance P, neuropeptide K), glucagon, gastrin-releasing peptide,VIP,PP, other biologically active peptides (ACTH, calcitonin, growth hormone-releasing hormone), prostaglandins, and bioactive amines (serotonin). These substances may or may not be released in sufficient amounts to cause symptoms. In patients with carcinoid tumors, elevated serum levels of PP were found in 43%, motilin in 14%, gastrin in 15%, and VIP in 6%. Foregut carcinoids are more likely to produce various GI peptides than midgut carcinoids. Ectopic ACTH production causing Cushing's syndrome is increasingly seen with foregut carcinoids (respiratory tract primarily), and in some series foregut carcinoid was the most common cause of the ectopic ACTH syndrome, accounting for 64% of all cases. Acromegaly due to GRF release occurs with foregut carcinoids, as does the somatostatinoma syndrome with duodenal carcinoids. The most common systemic syndrome is the carcinoid syndrome.

CARCINOID SYNDROME

Clinical Features The cardinal features at presentation as well as during the disease course are shown in Table 93-6. Flushing and diarrhea are the two most common symptoms, occurring in up to 73% initially and in up to 89% during the course of the disease. The characteristic flush is of sudden onset: it is a deep red or violaceous erythema of the upper body (especially the neck and face), often associated with a feeling of warmth, and occasionally associated with pruritus, lacrimation, diarrhea, or facial edema. Flushes may be precipitated by stress, alcohol, exercise, or certain foods such as cheese or by certain agents such as catecholamines, pentagastrin, and serotonin reuptake inhibitors. Flushing episodes may be brief, lasting 2 to 5 min, especially initially, or they may last for hours, especially later in the disease course. Flushing is usually seen with midgut carcinoids but can also occur with foregut carcinoids. With bronchial carcinoids the flushes are frequently prolonged for hours to days, reddish in color, and associated with salivation, lacrimation, diaphoresis, diarrhea, and hypotension. The flush associated with gastric carcinoids is also reddish in color but patchy in distribution over the face and neck. It may be provoked by food and have accompanying pruritus.

Diarrhea is present in 32 to 73% of patients initially and in 68 to 84% at some time in the disease course. Diarrhea usually occurs with flushing (85% of cases). The diarrhea is usually watery, with 60% having <1 L per day or diarrhea. Steatorrhea is present in 67%, and in 46% it is >15 g/day (normal <7 g). Abdominal pain may be present with the diarrhea or independently in 10 to 34% of cases.

Cardiac manifestations occur in 11% of patients initially and in 14 to 41% at some time in the disease course. The cardiac disease is due to fibrosis involving the endocardium, primarily on the right side, although left side lesions can occur also. The dense fibrous deposits are most common on the ventricular aspect of the tricuspid valve and less common on the pulmonary valve cusps. They can result in constriction of the valves and pulmonic stenosis is usually predominant, whereas the tricuspid valve is often fixed open, resulting in regurgitation. Up to 80% of patients with cardiac lesions develop heart failure. Lesions on the left side are much less extensive, are found in 30% at autopsy, and most frequently affect the mitral valve.

Other clinical manifestations include wheezing or asthma-like symptoms (8 to 18%) and pellagra-like skin lesions (2 to 25%). A variety of noncardiac problems due to increased fibrous tissue have been reported, including retroperitoneal fibrosis causing urethral obstruction, Peyronie's disease of the penis, intrabdominal fibrosis, and occlusion of the mesenteric arteries or veins.

Pathobiology In different studies covering 8876 patients with carcinoid tumors, carcinoid syndrome occurred in 8%, with a range of 1.4 to 18.4%. It occurs only when sufficient concentrations of tumor-secreted products reach the systemic circulation. In 91% of cases this occurs after metastasis to the liver. Rarely, the carcinoid syndrome can occur without hepatic metastases, caused by primary gut carcinoids with nodal metastases with extensive retroperitoneal invasion, pancreatic carcinoids with retroperitoneal lymph nodes, or carcinoids of the lung or ovary with direct access to the

systemic circulation. All carcinoid tumors do not have the same propensity to metastasize and cause the carcinoid syndrome (Table 93-3). Midgut carcinoids account for 60 to 67% of cases of carcinoid syndrome, foregut tumors for 2 to 33%, hindgut for 1 to 8%, and unknown primary sites for 2 to 15%.

One of the main secretory products of carcinoid tumors involved in the carcinoid syndrome is serotonin [5-hydroxytryptamine (5-HT)] (Fig. 93-1), which is synthesized from tryptophan. Up to 50% of dietary tryptophan can be used in this synthetic pathway by tumor cells, which can result in inadequate supplies for conversion to niacin; hence 2 to 5% of patients can develop pellagra-like lesions. Serotonin has numerous biologic effects including stimulating intestinal secretion, inhibiting absorption, stimulating increases in intestinal motility, and stimulating fibrogenesis. While 56 to 88% of carcinoid tumors are associated with serotonin overproduction, 12 to 26% of patients do not have the carcinoid syndrome. Serotonin overproduction is noted in 90 to 100% of patients with the carcinoid syndrome. Serotonin is thought to be predominantly responsible for the diarrhea by its effects on gut motility and intestinal secretion. Serotonin receptor antagonists (especially 5-HT_3antagonists) relieve the diarrhea in most patients. Prostaglandin E_2 and tachykinins may be important mediators of the diarrhea in some patients. Flushing is not relieved by serotonin receptor antagonists. In patients with gastric carcinoids the red, patchy pruritic flush is likely due to histamine release because it can be prevented by H_1 and H_2receptor antagonists. Numerous studies show tachykinins are stored in carcinoid tumors and released during flushing. Octreotide can relieve the flushing induced by pentagastrin in these patients without altering stimulated increase in plasma substance P, suggesting other mediators must be involved in the flushing. Both histamine and serotonin may be responsible for the wheezing as well as the fibrotic reactions involving the heart, causing Peyronie's disease and intraabdominal fibrosis. The exact mechanism of the heart disease is unclear. The valvular heart disease caused by the appetite-suppressant drug, dexfenfluramine, is histologically indistinguishable from that observed in carcinoid disease or after long exposure to 5-HT_2-selective ergot drugs. Metabolites of fenfluramine have high affinity for 5-HT_2receptors, whose activation is known to cause fibroblast mitogenesis. Lastly, high levels of 5-$HT_{2B}$and 5-HT_{2C}receptor transcripts are known to occur in heart valves. These observations support the conclusion that serotonin overproduction is important for the valvular changes, possibly by activating 5-HT_2receptors in the endocardium.

Patients may develop either a typical or atypical carcinoid syndrome (Fig. 93-1). In patients with the typical form, characteristically caused by a midgut carcinoid tumor, the conversion of tryptophan to5-HTP is the rate-limiting step. 5-HTP is rapidly converted to5-HT and stored in secretory granules of the tumor or in platelets. A small amount remains in plasma and is converted to5-HIAA, which appears in large amounts in the urine. These patients have an expanded serotonin pool size, increased blood and platelet serotonin levels, and increased urinary 5-HIAA. Some carcinoid tumors cause an atypical carcinoid syndrome thought to be due to a deficiency in the enzyme dopa decarboxylase; thus, 5-HTP cannot be converted to 5-HT (serotonin) and is secreted into the bloodstream. In these patients, plasma serotonin levels are normal but urinary levels may be increased because some 5-HTP is converted to 5-HT in the kidney. Characteristically, urinary 5-HTP and 5-HT are increased, but urinary 5-HIAA levels are only slightly elevated. Foregut carcinoids are the most likely to cause an atypical

carcinoid syndrome.

One of the most life-threatening complications of the carcinoid syndrome is the development of a carcinoid crisis. This is more frequent in patients who have intense symptoms from foregut tumors or have greatly increased urinary5-HIAAlevels (i.e., >200 mg/d). The crisis may occur spontaneously or be provoked by stress, anesthesia, chemotherapy, or a biopsy. Patients develop intense flushing, diarrhea, abdominal pain, and cardiac abnormalities including tachycardia, hypertension, or hypotension. If not adequately treated, it can be fatal.

Diagnosis The diagnosis of carcinoid syndrome relies on measurement of urinary or plasma serotonin or its metabolites in the urine. The measurement of5-HIAAis most frequently used. False-positive elevations may occur if the patient is eating serotonin-rich foods (e.g., bananas, pineapple, walnuts, pecans, avocados, or hickory nuts) or taking certain medications (e.g., cough syrup containing guaifanesin, acetaminophen, salicylates, or L-dopa). The normal range in daily urinary 5-HIAA excretion is between 2 and 8 mg. The 5-HIAA level has a 73% sensitivity and 100% specificity for carcinoid syndrome.

Most physicians use only the urinary5-HIAAexcretion rate; however, plasma and platelet serotonin levels, if available, may give additional information. Platelet serotonin levels are more sensitive than urinary 5-HIAA but are not generally available. If an atypical carcinoid syndrome is suspected and the urinary 5-HIAA is minimally elevated or normal, other urinary metabolites of tryptophan such as5-HTP or5-HTshould be measured.

Flushing occurs in a number of other conditions or diseases including systemic mastocytosis; chronic myelogenous leukemia with increased histamine release; menopause; reactions to alcohol or glutamate; and side effects of chlorpropamide, calcium channel blockers, and nicotinic acid. None of these conditions cause an increase in urinary5-HIAA.

The diagnosis of carcinoid tumor can be suggested by the carcinoid syndrome, by recurrent abdominal symptoms in a healthy-appearing individual, or by discovering hepatomegaly or hepatic metastases associated with minimal symptoms. Ileal carcinoids, which make up 25% of all clinically detected carcinoids, should be suspected in patients with bowel obstruction, abdominal pain, flushing, or diarrhea.

Serum chromogranin A levels are elevated in 50 to 100% of patients with carcinoid tumors, and the level correlates with tumor bulk. Serum chromogranin A levels are not specific for carcinoid tumors because they are also elevated in patients withPETs and otherNETs. Plasma neuron-specific enolase levels are also used as a marker of carcinoid tumors but are less sensitive than chromogranin A, being increased in only 17 to 47% of patients.

TREATMENT

Carcinoid Syndrome Treatment includes avoiding conditions that precipitate flushing, dietary supplementation with nicotinamide, treatment of heart failure with diuretics,

treatment of wheezing with oral bronchodilators, and controlling the diarrhea with antidiarrheal agents such as loperamide or diphenoxylate. If patients still have symptoms, serotonin receptor antagonists or somatostatin analogues are the drugs of choice.

There are 14 subclasses of serotonin (5-HT) receptors, and antagonists for most are not available. The $5\text{-HT}_1$and 5-HT_2receptor antagonists methysergide, cyproheptadine, and ketanserin have all been used to control the diarrhea but usually do not decrease flushing. The use of methysergide is limited because it can cause or enhance retroperitoneal fibrosis. Ketanserin diminishes diarrhea in 30 to 100% of patients. 5-HT_3receptor antagonists (ondansetron, tropisetron, alosetron) can control diarrhea and nausea in up to 100% of patients and occasionally ameliorate the flushing. A combination of histamine H_1 and H_2receptor antagonists (i.e., diphenhydramine and cimetidine or ranitidine) may control flushing in patients with foregut carcinoids.

Synthetic analogues of somatostatin (octreotide, lanreotide) are now the most widely used agents to control the symptoms of patients with carcinoid syndrome (Fig. 93-2). These drugs are effective at relieving symptoms and decreasing urinary5-HIAAlevels when self-administered every 6 to 12 h. Octreotide controls symptoms in>80% of patients, including the diarrhea and flushing, and 70% of patients show a >50% decrease in urinary 5-HIAA excretion. Patients with mild to moderate symptoms should initially be treated with 100 ug subcutaneously every 8 h. Individual responses vary, and patients have received doses as high as 3000 ug/d. About 40% of patients escape control after a median of 4 months, and the dose may need to be increased. Similar results are reported with lanreotide.

In patients with carcinoid crises, somatostatin analogues are effective at both treating the condition as well as preventing its development during known precipitating events such as surgery, anesthesia, chemotherapy, or stress. It is recommended that octreotide (150 to 250 ug subcutaneously every 6 to 8 h) be used 24 to 48 h before anesthesia and then continued throughout the procedure.

Sustained-release preparations of both octreotide [octreotide-LAR (long-acting release)] and lanreotide [lanreotide-PR (prolonged release)] are useful. Octreotide-LAR (30 mg/month) gives a plasma level ^{3}1 ng/mL for 25 days, whereas this level would require three to six injections per day of the non-sustained-release form. Lanreotide-PR is given intramuscularly every 10 to 14 days. Both sustained-release forms are highly effective.

Short-term side effects occur in 40 to 60% of patients receiving subcutaneous somatostatin analogues. Pain at the injection site and GI side effects (59% discomfort, 15% nausea, diarrhea) are the most common. They are usually short-lived and do not interrupt treatment. Important long-term side effects include gallstone formation, steatorrhea, and poor glucose tolerance. The overall incidence of gallstones/biliary sludge is 52%, with 7% of patients having symptomatic disease requiring surgical treatment.

Interferona is effective in controlling symptoms of the carcinoid syndrome, either alone or combined with hepatic artery embolization. The response rate is 42% for interferon a alone; when given with hepatic artery embolization, diarrhea was controlled for 1 year in

43% and flushing in 86% of patients.

Hepatic artery embolization alone or with chemotherapy (chemoembolization) has been used to control the symptoms of carcinoid syndrome. Embolization alone controls symptoms in up to 76% of patients, and chemoembolization (5-fluorouracil, doxorubicin, cisplatin, mitomycin) in 60 to 75% of patients. Hepatic artery embolization can have major side effects including nausea, vomiting, pain, and fever. In two studies, between 5 and 7% of patients died from complications of hepatic artery occlusion.

Other drugs have been used successfully in small numbers of patients to control the symptoms of carcinoid syndrome. Parachlorophenylalanine can inhibit tryptophan hydroxylase and the conversion of tryptophan to 5-HTP (Fig. 93-1). However, its severe side effects, including psychiatric disturbances, make it intolerable for long-term use.a-Methyldopa inhibits the conversion of 5-HTP to 5-HT; however, its effects are only partial.

Carcinoid Tumors (Nonmetastatic) Surgery is the only potentially curative therapy. Because the probability of metastases increases with increasing primary tumor size, the extent of surgical resection is determined accordingly. With appendiceal carcinoids, simple appendectomy is curative. With rectal carcinoids<1 cm, local resection is curative. With small-intestinal carcinoids <1 cm there is no consensus. Because 15 to 69% of small-intestinal carcinoids this size have metastases, some recommend a wide resection with *en bloc* resection of the adjacent lymph-bearing mesentery. If the carcinoid tumor is>2 cm for rectal, appendiceal, or small intestine, a full cancer operation should be done, including a right hemicolectomy for appendiceal carcinoid, an abdominoperineal or low anterior resection for rectal carcinoids, and an *en bloc* resection of adjacent lymph nodes for small-intestinal carcinoids. For carcinoids 1 to 2 cm in diameter in the appendix, a simple appendectomy is proposed by some, whereas others favor a right hemicolectomy. For 1- to 2-cm rectal carcinoids, a wide local full-thickness excision is recommended.

With type I or II gastric carcinoids, which are usually<1 cm, endoscopic removal is recommended. In type I or II gastric carcinoids if the tumor is >2 cm or if there is local invasion, some recommend total gastrectomy, others recommend antrectomy in type 1. For types I and II gastric carcinoids 1 to 2 cm, some recommend endoscopic treatment, others surgical treatment. With type III gastric carcinoids, if >2 cm, excision and regional lymph node clearance is recommended. Most tumors <1 cm are treated endoscopically.

PANCREATIC ENDOCRINE TUMORS

*Functional*PETsusually present with symptoms due to hormone excess. Only late in the course of the disease does the tumor itself cause prominent symptoms such as abdominal pain. In contrast, all of the symptoms due to *nonfunctional* PETs are due to the tumor. Thus, some functional PETs may present with severe symptoms with a small or undetectable primary tumor, whereas nonfunctional tumors almost always present late in their course when they are large and often metastatic. The mean delay between onset of continuous symptoms and diagnosis of a functional PET syndrome is 4 to 7 years. Therefore, the diagnoses are frequently missed for extended periods of time.

Treatment ofPETsrequires two different strategies. Treatment must be directed at the hormone excess state, such as the gastric acid hypersecretion in gastrinomas or hypoglycemia in insulinomas. Ectopic hormone secretion usually causes the presenting symptoms and can cause life-threatening complications. Except for insulinomas, >50% are malignant (Table 93-2); therefore treatment must also be directed against the tumor per se. These tumors are frequently not curable by surgery due to the extent of disease. Individual PETs are discussed below.

GASTRINOMA (ZOLLINGER-ELLISON SYNDROME) (See also Chap. 285)

A gastrinoma is aNETsecreting gastrin, a hormone that causes gastric acid hypersecretion (Zollinger-Ellison syndrome). The chronic hypergastrinemia results in marked gastric acid hypersecretion and growth of the gastric mucosa, with increased numbers of parietal cells and proliferation of gastricECLcells. The gastric acid hypersecretion characteristically causes peptic ulcer disease, often refractory and severe, as well as diarrhea. The most common presenting symptoms are abdominal pain (70 to 100%), diarrhea (37 to 73%), and gastroesophageal reflux disease (GERD) (30 to 35%) and 10 to 20% have diarrhea only. Although peptic ulcers may occur in unusual locations, most patients have a typical duodenal ulcer. The diagnosis of gastrinoma should be considered in patients with peptic ulcer disease with diarrhea; with peptic ulcers in an unusual or in multiple locations; and with peptic ulcer disease that is refractory to treatment or persistent, associated with prominent gastric folds, associated with findings suggestive ofMEN-1 (hyperparathyroidism, family history of ulcer or endocrinopathy, pituitary tumors), or without *Helicobacter pylori* present. *H. pylori* is present in >90% of patients with idiopathic peptic ulcers but is present in <50% of patients with gastrinomas. Chronic unexplained diarrhea should also suggest gastrinoma.

About 20 to 25% of patients haveMEN-1, and in most cases the hyperparathyroidism is present before the gastrinoma. These patients are treated differently from those without MEN-1; therefore, MEN-1 should be sought in all patients by family history and by measuring plasma calcium and plasma hormones (parathormone, growth hormone, prolactin).

Most gastrinomas (50 to 70%) are present in the duodenum, followed by the pancreas (20 to 40%) and other intraabdominal sites (mesentery, lymph nodes, biliary tract, liver, stomach, ovary). Gastrinomas may also occur in the left ventricular septum. InMEN-1 the gastrinomas are also usually in the duodenum (70 to 90%), followed by the pancreas (10 to 30%), and they are almost always multiple. Between 60 and 90% of gastrinomas are malignant (Table 93-2) with metastatic spread to lymph nodes and liver. Distant metastases to bone occur in 12 to 30% of patients with liver metastases.

Diagnosis The diagnosis of gastrinoma requires the demonstration of fasting hypergastrinemia and an increased basal gastric acid output (BAO; hyperchlorhydria). More than 98% of patients with gastrinomas have fasting hypergastrinemia, although in 40 to 60% the level may be less than 10 times normal. Therefore, when the diagnosis is suspected, a fasting gastrin level should be determined first. Gastric acid-suppressant drugs such as proton pump inhibitors (omeprazole, pantoprazole, lansoprazole) can suppress acid secretion sufficiently to cause hypergastrinemia and need to be

discontinued for a week before the gastrin determination. If the gastrin level is elevated, document that the gastric pH<2.5; hypergastrinemia secondary to achlorydria (atrophic gastritis, pernicious anemia) is one of the most common causes of hypergastrinemia. If the fasting gastrin is >1000 ng/L; 10 times normal) and the pH <2.5, which occurs in 40 to 60% of patients with gastrinoma, the diagnosis is established after ruling out the possibility of retained antrum syndrome by history. In patients with hypergastrinemia with fasting gastrins<1000 ng/L and gastric pH<2.5, other conditions such as *H. pylori* infections, antral G cell hyperplasia/hyperfunction, gastric outlet obstruction, or, rarely, renal failure can masquerade as a gastrinoma. To establish the diagnosis in this group, a determination of BAO and a secretin provocative test should be done. In >80%of patients with gastrinomas, BAO is elevated, i.e., 15 meq/h, and the secretin provocative test is positive, i.e., >200 ng/L increase in serum gastrin level.

TREATMENT

The gastric acid hypersecretion in patients with gastrinomas can be controlled in almost every case by oral gastric antisecretory drugs. Because of their long duration of action and potency, the proton pump inhibitors (H^+,K^+-ATPase inhibitors) are the drugs of choice. Histamine H_2-receptor antagonists are also effective, although more frequent dosing (every 4 to 8 h) and high doses are frequently required. In patients withMEN-1 with hyperparathyroidism, correction of the hyperparathyroidism increases the sensitivity to gastric antisecretory drugs and decreases the basal acid output.

With the increased ability to control acid hypersecretion, >50% of the patients who are not cured (>60% of patients) will die from tumor-related causes. At presentation, careful imaging studies are essential to localize the extent of the tumor (see below). About one-third of patients present with hepatic metastases; in <15% of those with hepatic metastases, the disease is limited so that surgical resection may be possible. Surgical cure is possible in 30% of all patients withoutMEN-1 or liver metastases (40% of all patients). In patients with MEN-1, long-term surgical cure is rare because the tumors are multiple, frequently with lymph node metastases. Therefore, all patients with gastrinomas without MEN-1 or a medical condition limiting life expectancy should undergo surgery.

INSULINOMAS

Insulinomas are endocrine tumors of the pancreas thought to be derived fromb cells that autonomously secrete insulin, which results in hypoglycemia. The average age of occurrence is in persons 40 to 50 years old. The most common clinical symptoms are due to the effect of the hypoglycemia on the central nervous system (neuroglycemic symptoms) and include confusion, headache, disorientation, visual difficulties, irrational behavior, or even coma (Chap. 334). Also, most patients have symptoms due to excess catecholamine release secondary to the hypoglycemia, including sweating, tremor, and palpitations. Characteristically these attacks are associated with fasting.

Insulinomas are generally small (>90% are <2 cm in diameter), usually solitary (90%), and only 5 to 15% are malignant. They almost invariably occur only in the pancreas, distributed equally in the pancreatic head, body and tail. Insulinomas should be suspected in all patients with hypoglycemia, especially with a history suggesting attacks

provoked by fasting or with a family history ofMEN-1. Insulin is synthesized as proinsulin, which consists of a 21-amino-acida chain and a 30-amino-acidb chain connected by a 33-amino-acid connecting peptide (C peptide). In insulinomas, in addition to elevated plasma insulin levels, elevated plasma proinsulin levels are found and C-peptide levels can be elevated.

Diagnosis The diagnosis of insulinoma requires the demonstration of an elevated plasma insulin level at the time of hypoglycemia. Other causes of fasting hypoglycemia include inadvertent or surreptitious use of insulin or oral hypoglycemic agents, severe liver disease, alcoholism, poor nutrition, or other extrapancreatic tumors. The most reliable test for diagnosing insulinoma is a fast up to 72 h with serum glucose, C-peptide, and insulin measurements every 4 to 8 h. If at any point the patient becomes symptomatic or glucose levels are persistently<2.2 mmol/L (40 mg/dL), the test should be terminated and repeat samples for the above studies obtained before glucose is given. Some 70 to 80% of patients will develop hypoglycemia during the first 24 h and 98% by 48 h. In nonobese normal subjects, serum insulin levels should decrease to >43 pmol/L (6 uU/mL) when blood glucose decreases to £2.2 mmol/L (40 mg/dL) and the ratio of insulin to glucose is <0.3 (in mg/dL). In addition to having an insulin level >6 uU/ml when blood glucose is £40 mg/dL, some investigators also require elevated C-peptide and serum proinsulin levels and/or insulin:glucose ratio >0.3 for the diagnosis of insulinoma. The effects of surreptitious use of insulin or hypoglycemic agents may be difficult to distinguish from the symptoms of insulinomas. The combination of proinsulin levels (normal in exogenous insulin/hypoglycemic agent users), C-peptide levels (low in exogenous insulin users), antibodies to insulin (positive in exogenous insulin users), and sulfonylurea levels in serum or plasma will allow the correct diagnosis to be made.

TREATMENT

Only 5 to 15% of insulinomas are malignant; therefore, after appropriate imaging, surgery should be performed. Between 75 and 95% of patients are cured by surgery. Before surgery the hypoglycemia can be controlled by frequent small meals and the use of diazoxide (150 to 800 mg/d). Diazoxide is a benozthiadiazide whose hyperglycemic effect is attributed to inhibition of insulin release; 50 to 60% of patients respond to diazoxide. Its side effects are sodium retention and GI symptoms such as nausea. Other agents effective in some patients to control the hypoglycemia include verapamil and diphenylhydantoin. Long-acting somatostatin analogues such as octreotide are acutely effective in 40% of patients. However, octreotide needs to be used with care because it inhibits growth hormone secretion and can lower plasma glucagon levels and so worsen the hypoglycemia.

For the 5 to 15% of patients with malignant insulinomas, the above drugs or somatostatin analogues are used initially. If they are not effective, hepatic arterial embolization, chemoembolization, or chemotherapy have been used. These will be discussed below.

GLUCAGONOMAS

Glucagonomas are endocrine tumors of the pancreas that secrete excessive amounts of glucagon that causes a distinct syndrome characterized by dermatitis, glucose

intolerance or diabetes, and weight loss. Glucagonomas mainly occur in persons between 45 and 70 years old. They are heralded clinically by a characteristic dermatitis (migratory necrolytic erythema; in 67 to 90%), accompanied by glucose intolerance (40 to 90%), weight loss (66 to 96%), anemia (33 to 85%), diarrhea (15 to 29%), and thromboembolism (11 to 24%). The characteristic rash usually starts as an annular erythema at intertriginous and periorificial sites, especially in the groin or buttock. It subsequently becomes raised and bullae form; when the bullae rupture, eroded areas form. The lesions can wax and wane. A characteristic laboratory finding is hypoaminoacidemia, which occurs in 26 to 100% of patients.

Glucagonomas are generally large tumors at diagnosis, with an average size of 5 to 10 cm. Between 50 and 80% occur in the pancreatic tail and 50 to 82% have evidence of metastatic spread at presentation, usually to the liver. Glucagonomas are rarely extrapancreatic and usually occur singly.

Diagnosis The diagnosis is confirmed by demonstrating an increased plasma glucagon level [normal is<150 ng/L]. In one study plasma glucagon levels were >1000 ng/L in 90%, between 500 and 1000 ng/L in 7%, and <500 ng/L in 3%. A plasma glucagon level>1000 ng/L is considered diagnostic. Other diseases causing increased plasma glucagon levels include renal failure, acute pancreatitis, hypercortisolism, hepatic failure, prolonged fasting, or familial hyperglucagonemia. Except for cirrhosis, these disorders do not usually increase plasma glucagon to >500 ng/L.

TREATMENT

Metastases are present at presentation in 50 to 80% of patients, so curative surgical resection is not possible. Surgical debulking in patients with advanced disease or other antitumor treatments may be beneficial (see below). Long-acting somatostatin analogues (octreotide, lanreotide) improve the skin rash in 75% of patients and may improve the weight loss, pain, and diarrhea but not the glucose intolerance.

SOMATOSTATINOMA SYNDROME

Somatostatinomas are endocrine tumors that secrete excessive amounts of somatostatin, which causes a syndrome characterized by diabetes melitus, gallbladder disease, diarrhea, and steatorrhea. The mean age of onset is 51 years. Somatostatinomas occur primarily in the pancreas and small intestine, and the frequency of the symptoms differs in each. The usual symptoms are more frequent in pancreatic than intestinal somatostatinomas: diabetes mellitus (95% vs. 21%), gallbladder disease (94% vs. 43%), diarrhea (92% vs. 38%), steatorrhea (83% vs. 12%), hypochlorhydria (86% vs. 12%), and weight loss (90% vs. 69%). Somatostatinomas occur in the pancreas in 56 to 74% of cases, with the primary location being in the pancreatic head. The tumors are usually solitary (90%) and large (mean diameter, 4.5 cm). Liver metastases are present in 69 to 84% of patients.

Somatostatin is a tetradecapeptide (Fig. 93-2), widely distributed in the central nervous system and gastrointestinal tract where it functions as a neurotransmitter or has paracrine and autocrine actions. It is a potent inhibitor of many processes, including release of almost all hormones, acid secretion, intestinal and pancreatic secretion, and

intestinal absorption. Most of the clinical manifestations are directly related to these inhibitory actions.

Diagnosis In most cases somatostatinomas have been found incidentally either at the time of cholecystectomy or during endoscopy. The presence of psammoma bodies in a duodenal tumor should particularly raise suspicion. Duodenal somatostatin-containing tumors are increasingly associated with von Recklinghausen's disease. Most of these do not cause the somatostatinoma syndrome as patients are usually asymptomatic and have normal plasma somatostatin levels. The diagnosis of somatostatinoma requires elevated plasma somatostatin levels.

TREATMENT

Pancreatic tumors are frequently metastatic at presentation (70 to 92%), whereas 30 to 69% of small-intestinal somatostatinomas have metastases. Symptoms are improved by octreotide treatment (Fig. 93-2).

VIPOMAS

VIPomas are endocrine tumors that secrete excessive amounts ofVIP, which causes a distinct syndrome characterized by large-volume diarrhea, hypokalemia, and dehydration. This syndrome is also called Verner-Morrison syndrome, pancreatic cholera, or WDHA syndrome (*w*atery *d*iarrhea, *h*ypokalemia, and *a*chlorhydria), which some patients develop. The mean age of patients is 49 years; however, the syndrome can occur in children; when it does, it is usually caused by a ganglioneuroma or ganglioneuroblastoma.

The principal symptoms are large-volume diarrhea (in 100%) severe enough to cause hypokalemia (80 to 100%), dehydration (83%), hypochlorhydria (54 to 76%), and flushing (20%). The diarrhea is secretory in nature, persists during fasting, and is almost always>1 L per day and >3 L per day in 70%. Most patients do not have accompanying steatorrhea (16%), and the increased stool volume is due to increased excretion of sodium and potassium, which, with the anions, accounts for the osmolality of the stool. Patients frequently have hyperglycemia (25 to 50%) and hypercalcemia (25 to 50%).

VIPis a 28-amino-acid peptide neurotransmitter, ubiquitously present in the central nervous system and GI tract. Its known actions include stimulation of small-intestinal chloride secretion as well as effects on smooth-muscle contractility, inhibition of acid secretion, and vasodilatory effects which explain most features of the clinical syndrome.

In adults 80 to 90% of VIPomas are pancreatic;VIP-secreting pheochromocytomas, intestinal carcinoids, and occasional ganglioneuromas account for the rest. These tumors are usually single; 50 to 75% are in the pancreatic tail and 37 to 68% have hepatic metastases at diagnosis.

Diagnosis The diagnosis requires the demonstration of an elevated plasmaVIPlevel and the presence of large-volume diarrhea. A stool volume of<700 mL/day excludes the diagnosis of VIPoma. A number of causes of diarrhea can be excluded by fasting the patient. Other diseases that can cause secretory large-volume diarrhea include

gastrinomas, chronic laxative abuse, carcinoid syndrome, systemic mastocytosis, diabetic diarrhea, AIDS, and rarely medullary thyroid cancer. Of these conditions, only VIPomas causes a marked increase in plasma VIP.

TREATMENT

The most important initial treatment in these patients is to correct their dehydration, hypokalemia, and electrolyte losses with fluid and electrolyte replacement. Patients may require 5 L/day of fluid and>350 mmol/day (350 meq/day) of potassium. Because 37 to 68% of adults withVIPomas have metastatic disease in the liver at presentation, a significant number of patients cannot be cured surgically. In these patients, long-acting somatostatin analogues such as octreotide or lanreotide (Fig. 93-2) are the drugs of choice.

Octreotide will control the diarrhea in 87% of patients. In nonresponsive patients, the combination of glucocorticoids and octreotide has proved helpful in a few. Other drugs that may be helpful include prednisone (60 to 100 mg/d), clonidine, indomethacin, phenothiazines, loperamide, lidamidine, lithium, proparanolol, and metoclopramide. Treatment of advanced disease with embolization, chemoembolization, and chemotherapy may also be helpful (see below).

NONFUNCTIONAL PANCREATIC ENDOCRINE TUMORS

NonfunctionalPETs are endocrine tumors that originate in the pancreas and either secrete no products or their secreted products do not cause a specific clinical syndrome. The symptoms are due entirely to the tumor per se. Nonfunctional PETs almost always secrete chromogranin A (90 to 100%), chromogranin B (90 to 100%),PP(58%), a-human chorionic gonadotropin (hCG) (40%), andb-HCG (20%), but none cause a specific syndrome. Patients with nonfunctional PETs usually present late in their disease course with invasive tumors and hepatic metastases (in 64 to 92%), and the tumors are usually large (72% >5 cm). These tumors are usually solitary except in patients withMEN-1, where they are multiple; they occur primarily in the pancreatic head; and though they do not cause a functional syndrome, they synthesize numerous peptides and cannot be distinguished from functional tumors by immunocytomchemistry.

The most common symptoms are abdominal pain (30 to 80%), jaundice (20 to 35%), and weight loss, fatigue, or bleeding; 10 to 15% are found incidentally. The average time from the beginning of symptoms to diagnosis is 5 years.

Diagnosis The diagnosis is established by histology in a patient with aPETwithout either clinical symptoms or elevated plasma hormone levels of one of the established syndromes (Table 93-2). Even though chromogranin A levels are elevated in almost every patient, this can be found in functional PETs, carcinoids, and other neuroendocrine disorders. PlasmaPP is increased in 22 to 71% of patients and should suggest the diagnosis in a patient with a pancreatic mass because it is usually normal in patients with pancreatic adenocarcinomas. However, elevated plasma PP is not diagnostic of this tumor because it is elevated in a number of other conditions such as chronic renal failure, old age, inflammatory conditions, and diabetes.

TREATMENT

Unfortunately, surgical curative resection can be considered in only a minority of patients because 64 to 92% present with metastatic disease. Treatment needs to be directed against the tumor itself using chemotherapy, embolization, chemoembolization, or hormonal therapy (see below).

GRFOMAS

GRFomas are endocrine tumors that secrete excessive amounts ofGRF that causes acromegaly. The frequency is not known. GRF (also called growth hormone-releasing hormone, GHRH) is a 44-amino-acid peptide, and 25 to 44% ofPETshave GRF immunoreactivity, although excess secretion is uncommon. GRFomas are lung tumors in 47 to 54% of cases, PETs in 29 to 30%, and small-intestinal carcinoids in 8 to 10% and up to 12% occur at other sites. Patients have a mean age of 38 years, and the symptoms are usually due to either acromegaly or the tumor per se. The acromegaly caused by GRFomas is indistinguishable from classic acromegaly. The pituitary abnormality is growth hormone-secreting somatotrope cell hyperplasia rather than a pituitary adenoma. The pancreatic tumors are usually large (>6 cm), and liver metastases are present in 39%. They should be suspected in any patient with acromegaly and an abdominal tumor, in a patient withMEN-1 with acromegaly, or in a patient without a pituitary adenoma with acromegaly or associated with hyperprolactinemia, which occurs in 70% of GRFomas. GRFomas are an uncommon cause of acromegaly. The diagnosis is established by performing plasma assays for GRF and growth hormone. The normal level for GRF is<5 ng/L (5 pg/mL) in men and<10 ng/L (10 pg/mL) in women. Most GRFomas have a plasma GRF level ³300 ng/L (300 pg/mL). Patients with GRFomas also have increased plasma insulin-like growth factor 1 levels similar to those in classic acromegaly. Surgery is the treatment of choice if diffuse metastases are not present. Long-acting somatostatin analogues such as octreotide or lanreotide (Fig. 93-2) induce responses in 75 to 100% of patients.

OTHER RARE PET SYNDROMES

Cushing's syndrome (ACTHoma) due to aPEToccurs in 4 to 16% of all patients with ectopic Cushing's syndrome. It occurs in 5% of cases of sporadic gastrinomas, almost invariably in patients with hepatic metastases, and is an independent, poor prognostic factor. Paraneoplastic hypercalcemia due to PETs releasing parathyroid hormone-related peptide is rare. The tumors are usually large, and liver metastases are usually present. PETs secreting calcitonin may cause a specific clinical syndrome. In one study, half the patients had diarrhea, which disappeared with resection of the tumor. InTable 93-2, this is a possible specific disorder because so few cases have been described.

TUMOR LOCALIZATION

Localization of the primary tumor and defining the extent of the disease are essential to the proper management of all carcinoids andPETs. Numerous tumor localization methods are used in both types ofNETs, including conventional imaging studies

[CTscanning, magnetic resonance imaging (MRI), transabdominal ultrasound, selective angiography] and somatostatin receptor scintigraphy (SRS). In PETs, endoscopic ultrasound (EUS) and functional localization by measuring venous hormonal gradients are also reported useful. Bronchial carcinoids are usually detected by a standard chest radiography and assessed by CT. Rectal, duodenal, colonic, and gastric carcinoids are usually detected by GI endoscopy.

PETsas well as carcinoid tumors possess high-affinity somatostatin receptors in both their primary tumors and their metastases. Of the five types of somatostatin receptors (sst_{1-5}), radiolabeled octreotide binds with high affinity to sst_2, lower for sst_3 and sst_5, and has very low affinity for sst_1 and sst_4. Between 90 and 100% of carcinoid tumors and PETs express sst_2, and many also have the other four sst subtypes. Interaction with these receptors can be used to localizeNETsusing [^{111}In-DTPA-D-Phe$_1$] octreotide (Fig. 93-2) and radionuclide scanning (SRS) as well as for treatment of the hormone excess state with octreotide or lanreotide. Because of its greater sensitivity than conventional imaging and its ability to localize tumor throughout the body at one time, SRS is now the imaging modality of choice for localizing both primary and metastastic NET tumors. SRS localizes tumors in 73 to 89% of patients with carcinoids and in 56 to 100% of patients with PETs, except for insulinomas. Insulinomas are usually small and have low densities of sst receptors, which results in SRS being positive in only 12 to 50% of patients with insulinomas.Figure 93-3 shows an example of the increased sensitivity of SRS in a patient with a gastrinoma. TheCT scan (Fig. 93-3, *top*) did not show any disease after primary tumor resection; however, hypergastrinemia remained, and the SRS demonstrated a metastasis in the liver (Fig. 93-3, *bottom*). Occasional false-positive responses with SRS can occur (12% in one study) because numerous other normal and abnormal cells can have high densities of sst receptors including granulomas (sarcoid, tuberculosis, etc.), thyroid diseases (goiter, thyroiditis), and activated lymphocytes (lymphomas, wound infections). For PETs located in the pancreas,EUS is highly sensitive, localizing 77 to 93% of insulinomas, which occur almost exclusively within the pancreas. EUS is less sensitive for extrapancreatic tumors. If liver metastases are identified by SRS, a CT scan orMRI is then recommended to assess the size and exact location of the metastases, because SRS does not provide reliable information on tumor size. Functional localization measuring hormone gradients after intraarterial calcium injections in insulinomas (insulin) or gastrin gradients after secretin injections in gastrinoma will be positive in 80 to 100% of patients. However, this method gives only regional localization and therefore is reserved for cases where other imaging tests are negative.

TREATMENT

Advanced Disease (Diffuse Metastatic Disease) The single most important prognostic factor for survival is the presence of liver metastases (Fig. 93-4). For patients with carcinoids without hepatic metastases, the 5-year survival is 80%; with limited liver metastases, it is also 80%; but with diffuse metastases, it is 50% (Fig. 93-4, *bottom*). With gastrinomas, the 5-year survival without liver metastases is 98%; with limited metastases in one hepatic lobe it is 78%; and with diffuse metastases, 16% (Fig. 93-4, *top*). A number of different modalities are effective, including cytoreductive surgery (removal of all visible tumor), treatment with chemotherapy, somatostatin analogues, interferona, hepatic embolization alone or with chemotherapy (chemoembolization),

radiotherapy, and liver transplantation.

Specific Antitumor Treatments Cytoreductive surgery is only possible in the 9 to 22% of patients who have limited hepatic metastases. No randomized studies have proven it extends life, but it appears to increase survival and therefore is recommended if possible.

Chemotherapy for metastatic carcinoid tumors has been disappointing, with response rates of 0 to 40% with various two- or three-drug combinations. Chemotherapy forPETs has been more successful, with tumor shrinkage reported in 30 to 70% of patients. The current regimen of choice is streptozotocin and doxorubicin.

Long-acting somatostatin analogues (octreotide, lanreotide) and interferon a rarely decrease tumor size (i.e., 0 to 17%); however, these drugs have tumoristatic effects, stopping additional growth in 50 to 95% of patients withNETs. How long tumor stabilization lasts or whether it prolongs survival has not been established.

Hepatic embolization and chemoembolization (with dacarbazine, cisplatin, doxorubicin, 5-fluorouracil, or streptozotocin) decrease tumor bulk and help control the symptoms of hormone excess. These modalities are generally reserved for patients in whom treatment with somatostatin analogues, interferon (carcinoids), or chemotherapy (PETs) fails.

Radiotherapy is being used with two different somatostatin radionuclides coupled by a DOTA-chelating group to octreotide (Fig. 93-2): [^{111}In-DTPA-D-Phe$_1$] octreotide (emits g rays, internal conversion, and Auger electrons) and yttrium-90 (emits high energy bparticles). The^{111}In compound showed disease stabilization in 40% and a decrease in tumor size in 30% of patients with advanced metastatic disease.

The use of liver transplantation has been abandoned for treatment of most metastatic tumors to the liver. However, for metastaticNETs it is still a consideration. Liver transplantation in 103 cases of malignant NETs (48 werePETs, 43 were carcinoids) achieved 2- and 5-year survival rates of 60% and 47%, respectively. However, recurrence-free survival was low (<24%). Liver transplantation may be justified for younger patients with metastatic NETs limited to the liver.

(Bibliography omitted in Palm version)

Back to Table of Contents

94. BLADDER AND RENAL CELL CARCINOMAS - *Howard I. Scher*, *Robert J. Motzer*

BLADDER CANCER

A transitional cell epithelium lines the urinary tract from the renal pelvis to the ureter, urinary bladder, and the proximal two-thirds of the urethra. Carcinomas may occur at any point, but generally 90% develop in the bladder, 8% in the renal pelvis, and 2% in the ureter or urethra. Overall, urinary bladder cancer is the fourth most common cancer in men and the ninth in women, with an estimated 53,200 new cases (38,300 males and 14,900 females) and 12,200 deaths (8100 males and 4100 females) predicted for the year 2000. The median age at diagnosis is 65 years. Once diagnosed, these tumors exhibit the tendency to recur over time and in new locations in the urothelial tract. As long as urothelium is present, continuous monitoring of the urothelial tract is required.

EPIDEMIOLOGY

Cigarette smoking is believed to contribute to up to 50% of the diagnosed urothelial cancers in men. The risk of developing a urothelial cancer in smokers is increased two- to fourfold relative to nonsmoking males and may persist for 10 years or longer after smoking is stopped. Other agents that have been implicated include exposure to aniline dyes, the drugs phenacetin and chlornaphazin, and external beam radiation. Chronic cyclophosphamide exposure increases risk nine-fold. Diets rich in meat and fat predispose to bladder cancer; ingestion of vitamin A supplements appears to be protective. Exposure to *Schistosoma haematobium*, a parasite found in many developing countries, is associated with an increase in both squamous (70%) and transitional cell (30%) carcinomas of the bladder.

PATHOLOGY

In the United States, 90 to 95% of bladder tumors diagnosed are transitional cell tumors. Pure squamous tumors with keratinization comprise 3%, adenocarcinomas 2%, and small cell tumors (with paraneoplastic syndromes)<1%. Adenocarcinomas develop primarily in the urachal remnant in the dome of the bladder or in the periurethral tissues. Some assume a signet cell histology. Lymphomas or melanomas are rare. Overall, 75% of tumors present as superficial lesions, 20% with muscle invasion, and 5% with metastatic disease. Of the transitional cell tumors, low-grade papillary lesions that grow on a central stalk are most common. They are very friable, have a tendency to bleed, and are at a high risk for recurrence, yet they rarely progress to the more lethal invasive variety. In contrast, carcinoma in situ (CIS) is a high-grade tumor that is considered a precursor of the more lethal muscle-infiltrating cancers. Tumors are rated by histologic type and grade. Grade I lesions (highly differentiated tumors) rarely progress to a higher stage, while grade III tumors usually do.

PATHOGENESIS

The multicentric nature of the disease and high rate of recurrence has led to the hypothesis that a field defect develops in the urothelium. Molecular genetic analyses of bladder tumors representing defined stages and different grades have shown a series of

primary chromosomal aberrations associated with cancer *development*, and *secondary changes* associated with *progression* to a more advanced stage. Using paired bladder tumor and normal tissues from the same patient, 9q deletions are an early event in cancer development, while 3p and 5q deletions were more prevalent in invasive vs. superficial tumors. Deletions of 17p (*TP53* locus), 18q (the *DCC* gene locus), and the *RB* gene locus on chromosome 13q24 were seen only in invasive disease, while deletions of 3p and 11p occur in both superficial and invasive tumors. p53 overexpression correlates with a higher probability of progression to a more advanced stage and bladder cancer mortality for patients with Ta, Tis, T1 and muscle-infiltrating lesions. These factors have not been routinely used for clinical decision-making.

CLINICAL PRESENTATION, DIAGNOSIS, AND STAGING

Hematuria occurs in 80 to 90% of patients with exophytic tumors, while irritative symptoms are more common for patients with in situ disease. The bladder is the most common source of gross hematuria (40%), but benign cystitis (22%) is a more common cause than bladder cancer (15%) (Chap. 48). Microscopic hematuria is more commonly of prostatic origin (25%), while 2% of bladder cancers produce microscopic hematuria. The documentation of hematuria requires evaluation with a urinary cytology, visualization of the urothelial tract by sonography or an intravenous pyelogram (IVP), and cystoscopy. Screening of asymptomatic subjects for hematuria has been evaluated and shown to increase the frequency of tumor diagnosis at an early stage. Screening has not been shown to confer a survival benefit. Ureteral obstruction may result in flank pain or discomfort. Symptoms of metastatic disease are documented less commonly as the first presenting sign of a urothelial cancer.

The endoscopic evaluation includes an examination under anesthesia to determine whether or not a palpable mass is present. A flexible endoscope is then inserted into the bladder, and a bladder barbotage is performed to assess the presence or absence of malignant cells. A visual inspection is then carried out and a cystoscopic map completed that includes the size, location, and number of lesions and their growth pattern (solid vs. papillary) (Fig. 94-1). An attempt should be made to resect all visible tumors along with a sample of the underlying muscle so that the depth of invasion can be documented. A notation is also made as to whether or not a tumor was completely removed. Random biopsies of "normal" mucosal areas are conducted to assess for a field defect. Each site that is biopsied should be recorded separately. For patients with a positive cytology and no apparent tumor within the bladder, selective catheterization of the ureter is required with retrograde examination to evaluate for upper tract disease.

The critical issue in management is whether or not the tumor has invaded muscle, which is difficult to assess with noninvasive procedures alone. Ultrasonography, computed tomography (CT) and/or magnetic resonance imaging (MRI) may assist in distinguishing a tumor that extends to the perivesical fat (T3b) from one that does not (T3a), and to document whether or not regional lymph nodes are involved (N+). They are also important in the assessment of the upper tracts. Distal metastases are assessed by CT of the abdomen, pulmonary x-rays, or radionculide imaging of the skeleton. The need for these studies is based in part on the local extent of the lesion.

The revised 1997 TNM (tumor, nodes, metastasis) staging system is illustrated inFig.

94-2. Ta lesions grow as exophytic lesions, whileCISstarts on the surface and tends to invade muscle. As the degree of muscle infiltration increases, the probability of nodal and subsequent distal spread also increases.

TREATMENT

Treatments are based on the extent and depth of invasion of the tumor within the primary site and the presence or probability of metastatic spread. At a minimum, the management of a tumor that has not invaded the bladder wall is a complete endoscopic resection with or without intravesical therapy. Recurrences are seen in 50% or more of cases, of which 5 to 20% will progress to a more advanced stage. The decision to recommend additional therapy is based on the histologic subtype, the number of lesions, the depth of invasion, and whether or notCIS is present. Solitary papillary lesions are generally treated by surgery alone. Intravesical therapy is usually recommended for recurrent disease.

CISfrequently follows a more aggressive course. As such, intravesical therapy is generally recommended earlier in the clinical course. The standard treatment for a tumor that has invaded muscle, either at the time of diagnosis or following treatment for superficial disease, is radical cystectomy. Depending on the findings at surgery, systemic chemotherapy may or may not be advised.

Superficial Disease Intravesical therapies are applied in two contexts: as an adjuvant to a complete endoscopic resection to prevent recurrence, or, less commonly, to eliminate disease that cannot be controlled by endoscopic resection alone. Intravesical treatments are advised for patients with four or more recurrences in a given year, >40% involvement of the bladder surface by tumor, the presence of diffuseCIS, or documented T1 disease. A number of agents are available, but based on randomized comparisons, Bacillus Calmette-Guerin (BCG) is considered standard. Thiotepa, doxorubicin, mitomycin-C, and interferon have also been used. Side effects include dysuria, frequency, and, depending on the drug, myelosuppression or a contact dermatitis (from mitomycin C). Rarely, intravesical BCG may produce a systemic illness associated with granulomatous infections in multiple sites that requires anti-tuberculin therapy for control. Significant BCG toxicities occur in <6% of patients.

Following endoscopic resection, patients are reevaluated at 3-month intervals to ensure that no recurrences have developed. Those with persistent disease or new tumors are generally considered for a second course ofBCG or intravesical chemotherapy. Those with persistent disease may be considered for cystectomy, although the specific indications vary. Obvious candidates are those with new invasive tumors or persistentCIS, and those with bladder function that has been compromised to the point where persistent pain, blood loss, frequency, or a severely limited bladder capacity is present. Recurrences may develop anywhere along the urothelial tract, including the renal pelvis, ureter, or urethra. In fact, one consequence of the "successful" treatment of tumors in the bladder is an increase in the frequency of extravesical recurrences.

Muscle-Infiltrating Disease The treatment of a tumor that has invaded muscle can be separated into control of the primary tumor and control of systemic disease. Radical cystectomy is considered the standard treatment, although in selected cases

bladder-sparing approaches using an aggressive endoscopic resection, partial cystectomy, or a combined modality approach with resection, systemic chemotherapy, and external beam radiation therapy are used. The latter should not be considered outside of a research setting.

Radical cystectomy involves an evaluation of the pelvic lymph nodes, removal of the primary tumor, and creation of a conduit or reservoir for urinary flow. At the time of surgery, grossly abnormal lymph nodes are evaluated by frozen section. If metastases are confirmed, the procedure is often aborted unless a diversion is required for palliation of local symptoms. The results of treatment of node-positive disease are shown in Table 94-1. In males radical cystectomy involves the removal of the bladder, prostate, seminal vesicles, proximal vas deferens, and proximal urethra, with a margin of adipose tissue and peritoneum. Impotence is universal unless the *nervi erigentes*, responsible for erectile capacity, can be preserved. In females the procedure includes removal of the bladder, urethra, uterus, fallopian tubes, ovaries, anterior vaginal wall, and surrounding fascia.

Urine flow is directed through either an internal reservoir that drains to the urethra or the abdominal wall, or via a Bricker procedure in which urine flows through an ileal conduit from the ureters to the abdominal wall, where it is collected in an external appliance without an internal reservoir. A segment of colon, jejunum, or ileum can be used to bridge the gap between the ureters and the skin. Use of absorbable sutures may prevent formation of calculi at the sutures. A uretero-ileal conduit probably is the most widely used. A syndrome characterized by hypochloremic acidosis, hyperkalemia, hyponatremia, and uremia has been described when a segment of jejunum is utilized. Concurrent diseases in the bowel, such as ulcerative colitis or Crohn's disease, may hinder the use of resected bowel.

Alternatives to an external appliance include internal reservoirs that are created from detubularized bowel segments and are periodically self-catheterized by the patient. A number of procedures have been described that use either ileocecal or ileal reservoirs, which are anastamosed to either the abdominal wall or the urethra. When an anastomosis to the urethra is created, primarily in men with no urethral disease, the patient can then void in a manner that is similar to natural voiding. Several indications for urethrectomy (includingCIS or exophytic tumor in the urethra and diffuse CIS in the urinary bladder) preclude the creation of a urethral anastomosis. Continent reservoirs are being applied with increasing frequency, but are still not constructed for the majority of patients, for technical or disease-related reasons. Intercurrent diseases, impaired renal function, hesitancy to prolong the surgical trauma, dilated ureters, and bowel diseases all decrease the use of continent reservoirs. Patients with ureterosigmoid diversion require periodic colonoscopy because of the risk of cancer.

Cystectomy is major surgery, and appropriate medical clearance is essential. This includes optimizing cardiac medications and nutritional status. In approximately 5 to 10% of cases, depending on the location of the tumor, a partial cystectomy is possible. This procedure can be considered when a lesion develops on the dome of the bladder where a 2-cm margin can be achieved,CIS is absent in other sites of the bladder, and bladder capacity is adequate after the tumor is removed. Carcinomas in the ureter or in the renal pelvis are treated by nephroureterectomy with a bladder cuff.

Indications for cystectomy include: (1) muscle-invading tumors not suitable for segmental resection; (2) low-stage tumors unsuitable for conservative management due to, for example, multicentric and frequent recurrences resistant to intravesical instillations; (3) high grade tumors (T1G3) associated withCIS or bladder symptoms such as frequency or hemorrhage rendering the patient a "bladder cripple." Outcomes are reported on the basis of 5-year survivals. As shown in Table 94-2, survival varies inversely with depth of invasion and lymph node status. For the majority of cases, however, extension to a single lymph node predicts a poor outcome with a median time to recurrence of 22 months. In some countries external beam radiation therapy is considered standard. This is not the case in the United States, where its role is limited to those patients deemed unfit for cystectomy or those with unresectable local disease, and as part of an experimental approach that seeks to spare the bladder.

Metastatic Disease Patients with metastatic disease include those whose tumor has recurred after definitive local treatment and those who present with metastases. A number of chemotherapeutic agents have shown activity as single agents, of which cisplatin, paclitaxel, and gemcitabine are considered most active (Table 94-3). Responses to single agents are generally incomplete and not durable. Using multidrug regimens, response rates in excess of 50% have been reported with combinations such as M-VAC (methotrexate, vinblastine, doxorubicin, and cisplatin), PT (cisplatin and paclitaxel), and gemcitabine variants. Based on randomized comparisons, M-VAC is considered standard but can be associated with significant toxicities, including neutropenia and fever; mucositis in 10 to 20%; a decrease in renal and auditory function; and a peripheral neuropathy. Alopecia is universal; fatigue can be dose-limiting in some cases. More recently 2- and 3-drug combinations based on cisplatin/carboplatin, paclitaxel, and gemcitabine have been explored. In a direct comparison to M-VAC, gemcitabine/cisplatin showed similar response proportions and survival with fewer side effects. Long-term survival may be obtained in 10 to 15% of patients with metastatic disease and 20 to 25% of patients with unresectable nodal disease at presentation. In general, the proportion of patients rendered tumor-free is higher in patients with disease limited to nodal sites as opposed to visceral or bone sites. Patients with adverse features, such as a compromised performance status, visceral disease, or bone metastases, are rarely cured with chemotherapy alone. In these cases, median survivals rarely exceed 6 months.

Chemotherapy for Invasive Disease Chemotherapy can be given before (neoadjuvant) or after (adjuvant) definitive local therapy. Cumulative results of nonrandomized phase II trials have shown that the proportion of bladders rendered free of tumor varies inversely with T stage; but only 20 to 25% of bladders are tumor-free after chemotherapy alone. To date, neoadjuvant chemotherapy has not been shown to prolong life. Several groups are investigating bladder-sparing strategies but these approaches are not considered routine practice. The need for adjuvant therapy is based on a pathologic determination of risk. In general, the finding of nodal disease at surgery, extravesical tumor extension, or vascular invasion in the resected specimen are considered indications for postoperative adjuvant therapy. When administered, a minimum of four cycles at full dose is recommended.

Overview Superficial TaG1 lesions rarelyprogress to an invasive lesion and can be

handled with an endoscopic resection. Muscle-invasive disease may require both aggressive local therapy and systemic therapy of micrometastases for cure, while metastatic urinary bladder cancer is the most lethal for the majority of patients. Only a small proportion of patients with metastatic disease can be cured with chemotherapy. Current refinements in therapy include identifying subgroups of patients with superficial disease where the intensity of follow-up can be reduced or where intravesical therapy is needed. For muscle-invasive disease, efforts are being made to identify patients for whom organ preservation is possible without compromising overall survival, as well as those with subclinical micrometastases for whom systemic therapy is needed for cure. Efforts to improve therapy include better surgical techniques and the incorporation of newly identified chemotherapeutic agents into combination regimens. For the majority of patients, combined modality approaches are essential to optimal management.

RENAL CELL CARCINOMA

Renal cell carcinoma accounts for 90 to 95% of malignant neoplasms arising from the kidney. Notable features include refractoriness to cytotoxic agents, infrequent but reproducible responses to biologic response modifiers such as interferona and interleukin (IL) 2, and a variable clinical course for patients with metastatic disease, including anecdotal reports of spontaneous regression.

EPIDEMIOLOGY AND ETIOLOGY

In the year 2000, 31,200 new cases of renal cancer were diagnosed, and 11,900 people died of the disease. The male:female ratio is 2:1. Incidence peaks between the ages of 50 and 70, although this malignancy may be diagnosed at any age. Many environmental factors have been investigated as possible contributing causes. The strongest association is with cigarette smoking (accounting for 20 to 30% of cases) and obesity. The risk is increased for patients who have acquired cystic disease of the kidney associated with end-stage renal disease.

Most cases are sporadic, although familial forms have been reported. One is associated with von Hippel-Lindau (VHL) syndrome. Nearly 35% of patients with VHL disease develop renal cell cancer. An increased incidence has also been reported for patients with tuberous sclerosis and polycystic kidney disease.

Most of the cancers arise from the epithelial cells of the proximal tubules. A number of genetic alterations have been described, of which abnormalities on chromosome 3 are most frequent. A t(3;8) translocation was first described in a pedigree of patients with the familial form of the disease, while deletions of 3p21-26 (where *VHL* maps) have been identified in familial as well as sporadic tumors. *VHL* mutations are identified in a high proportion of sporadic, nonpapillary renal cell cancers and associated cell lines.

PATHOLOGY

Renal cell neoplasia represents a heterogeneous group of tumors with distinct histopathologic, genetic, and clinical features ranging from benign to high-grade malignant. Categories include clear cell carcinoma (60% of cases), papillary (5 to 15%), chromophobic tumors (5 to 10%), oncocytomas (5 to 10%), and collecting or Bellini duct

tumors (<1%). Clear cell tumors are characterized by tumor cells with clear cytoplasm and consistently show a deletion of 3p. Papillary tumors tend to be bilateral and multifocal. Trisomy 7 and/or 17 are the most frequent genetic markers. Chromophobic tumors are characterized by multiple chromosomal losses but do not exhibit 3p deletions; they also have a more indolent clinical course. Oncocytomas have a characteristic morphology including a deeply eosinophilic cytoplasm, do not exhibit 3p deletions or trisomy 7 or 17, and are considered benign neoplasms. In contrast, Bellini duct carcinomas are very rare and are thought to arise from the collecting ducts within the renal medulla. They tend to afflict younger patients and are very aggressive tumors.

CLINICAL PRESENTATION

The presenting signs and symptoms include hematuria, abdominal pain, and a flank or abdominal mass. This classic triad occurs in 10 to 20% of patients. Other symptoms are fever, weight loss, anemia, and a varicocele (Table 94-4); the tumor can be found incidentally on a radiograph.

The presentation has changed over the past two decades, due to the advent and widespread use of radiologic cross-sectional imaging procedures (CT, ultrasound,MRI). The more frequent use of sensitive abdominal imaging modalities in recent years contributes to earlier detection, including incidental low-stage renal masses detected during evaluation for other medical conditions. The increasing number of incidentally discovered low-stage tumors contributes to an improved 5-year survival for patients with renal cell carcinoma and increased use of nephron-sparing surgery (partial nephrectomy).

A spectrum of paraneoplastic syndromes has been associated with these malignancies, including erythrocytosis, hypercalcemia, nonmetastatic hepatic dysfunction (Stauffers' syndrome) and acquired dysfibrinogenemia. Erythrocytosis is present at presentation in only about 3% of patients. More frequently anemia, a sign of advanced disease, is reported.

The standard evaluation of patients with suspected renal cell tumors includes aCT scan of the abdomen and pelvis, a chest radiograph, urine analysis, and urine cytology. A CT of the chest is warranted if metastatic disease is suspected from the chest radiograph, as it will detect significantly smaller lesions, and their presence may influence the approach to the primary tumor.MRI is useful in evaluating the inferior vena cava in cases of suspected tumor involvement or invasion by thrombus, as well as for patients in whom iodinated contrast cannot be administered owing to either allergy or renal dysfunction. In clinical practice any solid renal masses should be considered malignant until proved otherwise and require a definitive diagnosis. If no metastases are demonstrated, surgery is indicated, even if there is invasion of the renal vein. The differential diagnosis of a renal mass includes cysts, benign neoplasms (adenoma, angiomyolipoma, oncocytoma), inflammatory lesions (pyelonephritis or abscesses), and other primary or metastatic malignant neoplasms. Other malignancies that may involve the kidney include transitional cell carcinomas of the renal pelvis, sarcoma, lymphoma, Wilms' tumor, and metastatic disease, especially from melanoma primaries. All of these are less common than renal cell carcinoma as kidney masses.

STAGING AND PROGNOSIS

Two staging systems used commonly are the Robson classification and the American Joint Committee on Cancer (AJCC) staging system. According to the former, stage I tumors are confined to the kidney; stage II tumors extend through the renal capsule but are confined to Gerota's fascia; stage III tumors involve the renal vein or vena cava (stage III A) or the hilar lymph nodes (stage III B); and stage IV disease includes tumors that are locally invasive to adjacent organs (excluding the adrenal gland) or distant metastases. Five-year survival rate varies by stage: 66% for stage I, 64% for stage II, 42% for stage III, and 11% for stage IV. The prognosis for patients with stage IIIA lesions is similar to that of stage II disease, whereas the 5-year survival rate for patients with stage IIIB lesions is only 20%, closer to that of stage IV.

TREATMENT

Localized Tumors The standard management for stage I or II tumors and selected cases of stage III disease is radical nephrectomy. This procedure involves en bloc removal of Gerota's fascia; its contents including the kidney, the ipsilateral adrenal gland, and adjacent hilar lymph nodes. The role of a regional lymphadenectomy is controversial. For patients with stage IIIA disease, the tumor should be resected from the renal vein or vena cava.

In selected patients who have only one kidney, a partial nephrectomy may be performed, depending on the size and location of the lesion. Partial nephrectomy may also be performed for patients with bilateral tumors, accompanied by a radical nephrectomy on the opposite side. Partial nephrectomy techniques are being applied electively to resect small masses for patients with a normal contralateral kidney. There is no proven role for adjuvant chemotherapy, immunotherapy, or radiation therapy following successful surgical removal of the tumor, even in cases with a poor prognosis.

Advanced Disease Metastatic renal cell carcinoma, for which there is no effective therapy, is associated with dismal survival. A number of options have been explored, including hormonal therapy, chemotherapy (cytotoxic agents), and immunotherapy. Responses to hormonal therapy (progestins) are rare (1 to 2%) and of short duration. No chemotherapy agent has been shown to consistently produce tumor regressions in>20% of patients.

Two biologic therapies, interferona andIL-2, have been studied extensively for the treatment of advanced disease. Both reproducibly produce responses in 10 to 20% of patients; the response is durable in fewer than 5% of patients. It was the observation of occasional durable complete remissions that resulted in the Food and Drug Administration's approval of IL-2 as a treatment for this disease. IL-2 is usually administered by infusion of 720,000 IU/k every 8 h per day for 5 to 7 days. Toxicities from IL-2 include a capillary leakage syndrome, fever, chills, fatigue, and hypotension.

Surgery in the Setting of Metastases Nephrectomy may be indicated in highly selected cases for the alleviation of symptoms, including pain or recurrent urinary hemorrhage, and particularly if the latter is severe or associated with obstruction. Some physicians advocate the performance of a nephrectomy in the presence of metastases

in the hope either that a spontaneous regression will occur or that the sensitivity to a cytokine will be increased. The observed frequency of spontaneous regression, 0.8%, coupled with the morbidity and mortality of the procedure itself, does not justify the approach. Nephrectomy in the presence of metastatic disease followed by IFN-a is associated with a modest survival benfit over IFN-aalone.

There are reports of long-term survival at rates of 15 to 50% for patients who relapse following nephrectomy at a solitary site and undergo surgical resection of the metastasis. Because renal cell tumors are radioresistant, surgical resection is also advised for palliation of solitary central nervous system metastases, repair of actual or impending fractures in weight-bearing bones, or relief of spinal cord compression.

Observation Alone Renal cell carcinoma is one of several malignancies in which spontaneous regressions have been reported anecdotally. A more frequent occurrence is prolonged periods of stable disease: up to 10% of patients with metastatic disease show no progression for >12 months. Because responses to systemic therapy are uncommon, and all systemic therapies are associated with treatment-related toxicity, an option for management in asymptomatic patients with metastases is close observation until evidence of disease progression or symptoms occur, at which time appropriate therapy is initiated. It is important to document the presence of progressive disease before initiating an experimental treatment; this will avoid attributing stable disease to the drug when it may be a feature of the tumor.

CARCINOMA OF THE RENAL PELVIS AND URETER

About 500 cases of renal pelvis and ureter cancer occur each year; nearly all are transitional cell carcinomas similar to bladder cancer in biology and appearance. This tumor also is associated with chronic phenacetin abuse and with Balkan nephropathy, a chronic interstitial nephritis endemic in Bulgaria, Greece, Bosnia-Herzegovina, and Romania.

The most common symptom is painless gross hematuria, and the disease usually is detected onIVPduring the workup for hematuria. Patterns of spread are like those in bladder cancer. For disease localized to the renal pelvis and ureter, nephroureterectomy (including excision of the distal ureter with a portion of the bladder) is associated with a 5-year survival of 80 to 90% for low-grade lesions. More invasive or histologically poorly differentiated tumors are more likely to recur locally and metastasize. Metastatic disease is treated with M-VAC or CMV (cisplatin, methotrexate, vinblastine) chemotherapy, as used in bladder cancer, and the outcome is similar to that for metastatic transitional cell cancer of bladder origin.

(Bibliography omitted in Palm version)

Back to Table of Contents

95. HYPERPLASTIC AND MALIGNANT DISEASES OF THE PROSTATE - *Howard I. Scher*

The process of aging is associated with an increasing frequency of both benign and malignant alterations of the prostate gland. These conditions reflect the uncontrolled growth of both the stromal and epithelial components of the gland. Autopsies of men in the eighth decade of life show hyperplastic changes in>90% and malignant changes in>70%. Most men with benign or malignant conditions of the prostate are not diagnosed during their lifetimes. The high prevalence of these diseases, coupled with comorbid conditions and competing causes of death that are frequent in this age group, mandates a careful consideration of the risk/benefit ratio of any proposed intervention. Management is centered on the continual reassessment of the disease as it unfolds in the individual. For the benign proliferative disorders, the symptoms of urinary frequency, infection, and potential for obstruction are counterbalanced by the side effects and complications of medical or surgical therapy. For malignant disease, the risk of developing symptoms or death from cancer is balanced against treatment efficacy and treatment-related morbidity for interventions proposed at different points in the natural history.

The incidence and mortality of prostate cancer have declined over the past few years. The decline is not clearly related to any meaningful decrease in the disease or its severity. Instead, the number of cases diagnosed increased dramatically in the early 1990s based on the widespread use of serum prostate-specific antigen (PSA) levels. The test led to the diagnosis of more asymptomatic cancers, some of which may never have produced symptoms -- so-called lead-time bias (Chap. 80). Screening for prostate cancer has not been proven effective in prospective randomized trials. Prostate cancer is the most common cancer diagnosis and the second leading cause of cancer death in men. In 2000, 180,400 cases were diagnosed and 31,900 men died of prostate cancer, down from the peak of 352,000 new cases in 1996. The projected lifetime risk of developing prostate cancer for a 50-year-old man is 42%, of being diagnosed is 9.5%, and of dying from prostate cancer is 2.9%.

ANATOMY

The prostate is located in the pelvis and is surrounded by the rectum, bladder, dorsal and periprostatic venous complexes, musculature of the pelvic sidewall, the urethral sphincter (responsible for passive urinary control), the pelvic plexus, and cavernous nerves (which innervate the pelvic organs and corpora cavernosa). It is divided into a peripheral zone, a central zone, and a transition zone. The anterior surface is covered by the fibromuscular stroma. Most cancers develop in the peripheral zone, while nonmalignant proliferation occurs predominantly in the transition zone. The functional unit is the glandular acinus, which consists of an epithelial compartment including epithelial, basal, and neuroendocrine cells, and a stromal compartment including fibroblasts and smooth-muscle cells. These compartments are separated by a basement membrane.PSA and prostate-specific acid phosphatase are produced in the epithelial cells. Both stromal and epithelial cells express androgen receptors and depend on androgens for growth. Additional growth regulatory signals occur via paracrine signaling between the two compartments. In cancer, the relationship between stromal and epithelial elements contributes to growth both in the primary and in

metastatic sites. The major circulating androgen in the blood is testosterone, which is converted to dihydrotestosterone, the active form, by 5a-reductase. Changes in prostate size occur during two distinct periods: diffuse enlargement during puberty and in focal regions in the periurethral area after the age of 55.

DIAGNOSIS AND SCREENING

Symptoms Most cancers are asymptomatic in their early stages. By contrast, benign proliferative disorders may encroach on the urethra early in the clinical course, giving rise to symptoms of outlet obstruction such as hesitancy, intermittent voiding, diminished stream, incomplete emptying, and postvoid leakage. For the patient with symptoms, the history is focused on the urinary tract to identify other causes of voiding dysfunction. For quantification of symptoms, the preferred questionnaire is the self-administered American Urological Association (AUA) *Symptom Index* in which the symptoms can be classified as mild, moderate, or severe on the basis of seven questions (Table 95-1). This index is useful in planning and in follow-up. Over time, the resistance to the flow of urine reduces the compliance of the detrusor muscle, resulting in nocturia, urgency, and bladder instability and ultimately in urinary retention. The relationship between the signs and symptoms of obstruction and prostate size is not straightforward, and a small gland does not exclude significant blockage. In severe cases the bladder may be palpable on physical examination. Infection, tranquilizing drugs, antihistamines, or alcohol can precipitate urinary retention.

Obstructive symptoms are distinct from irritative symptoms such as frequency, dysuria, or urgency, which may occur from infectious, inflammatory, or neoplastic diseases. Conditions that can mimic cancer include acute prostatitis, granulomatous prostatitis, and prostate calculus. Prostatitis usually produces induration and/or pain and is treated with antibiotics. Prostate cancer may manifest in the same manner, and the distinction can only be established histologically, but a biopsy should not be performed before a trial of antibiotics if prostatitis is a possible diagnosis. In cases where the tumor has extended beyond the confines of the gland, symptoms of hematospermia or erectile dysfunction may occur. Prostate cancer may also present with pain secondary to bone metastases, although many patients are asymptomatic despite extensive spread. Less common presentations include myelophthisic disorders, disseminated intravascular coagulation, or spinal cord compression. The proportion of men diagnosed at these late stages has also decreased significantly as a result ofPSA-based detection strategies.

Physical Examination The standard evaluation for prostatic diseases includes the digital rectal examination (DRE). It should be performed with careful attention to the size and consistency of the gland, the presence of lesions within the gland, or evidence of extension beyond its confines. Its importance can not be overemphasized. The posterior surfaces of the lateral lobes, where carcinoma begins most often, are easily palpable on DRE. Carcinoma characteristically is hard, nodular, and irregular, but induration may also be due to fibrous areas in a benign hyperplastic background or to calculi. Extraprostatic extension to the seminal vesicles can often be detected by rectal examination, while scrotal and/or lower extremity lymphedema secondary to infiltration of pelvic lymph nodes indicates extensive disease. The need for establishing a histologic diagnosis is based on the finding of an abnormal DRE or an elevated serumPSAlevel.

Prostate-Specific Antigen PSA is a serine protease that is produced by both nonmalignant and malignant epithelial cells. In serum, it circulates as an inactive complex with two protease inhibitors -- a_1-antichymotrypsin andb_2-macroglobulin. PSA is prostate specific but not prostate cancer specific and is measured most commonly by radioimmunoassay. The normal range is 0 to 4 ng/mL; ~30% of men with a PSA in the range of 4 to 10 ng/mL and 50% of those with a PSA>10 ng/mL will have cancer. Between 20 and 25% of men with an abnormalDRE have cancer at biopsy, and 20% of men have cancer detected when the PSA is in the normal range. African American men normally have higher PSA levels, even if they do not have prostate cancer. They also have a 50% higher risk of prostate cancer. The reason for the racial differences is not known.

Several refinements have been proposed to increase the sensitivity of the test for younger men more likely to die of the disease, while reducing the frequency of diagnosing cancers of low biologic potential in elderly men more likely to die of other causes. These modifications include age-specific reference ranges, using a lower "upper" limit of normal for younger males and higher "upper" limit for older individuals. Prostate-specific antigen density (PSAD) is calculated by dividing the serumPSAlevel by the estimated prostate weight calculated from transrectal ultrasonography (TRUS). It was proposed to correct for the influence of benign prostatic hyperplasia (BPH) on the measured level of PSA. Values<0.10 are consistent with BPH, while values>0.15 suggest the presence of cancer. PSAD levels also increase with age.

PSA velocity is derived from calculations of the rate of change inPSAbefore the diagnosis of cancer was established. Increases of>0.75 ng/mL per year are suggestive of cancer. For a 50-year-old male, an increase from 2.5 to 3.9 ng/mL in a 1-year period would warrant further testing, even though the level is still in the "normal" range. Free and complexed PSA measurements are used to determine which men require a biopsy when the PSA level is in the range of 4 to 10 ng/mL. In cancer, the level of free PSA is lower. Using a 25% threshold of free PSA for patients with levels in the range of 4 to 10 ng/mL, specificity was improved by 20% while maintaining a sensitivity of 95%. Further refinements to increase the specificity of distinguishing benign and malignant conditions involve the determination of the ratios of free to total, complexed to total, and free to complexed PSA. Using normal ranges for free/total PSA of >0.15, for complexed/total PSA of <0.70, and for free/complexed PSA of >0.25 improved specificity in one study by 20%. These modifications are designed to reduce the frequency of biopsies in men without cancer. Figure 95-1 illustrates a diagnostic algorithm based on the DREand PSA findings.

Transrectal Ultrasonography Most cancers are hypoechoic by ultrasonography. Unfortunately, no single finding on ultrasound permits univeral distinction between cancer and benign conditions and identification of extracapsular disease. Cancers <5 to 7 mm, those that are well differentiated, and those located in the transition zone are difficult to distinguish from the normal prostate. The primary role of ultrasound is to ensure accurate sampling of any index lesions and of the gland during biopsy. Routine sampling includes a minimum of six cores from the peripheral zone of the gland, each of which is identified and labeled separately for histologic examination. A biopsy session, defined as one that procures four or more cores from widely separated areas of the

prostate, has a sensitivity of ~80% for the detection of a cancer.TRUS has also been used in the assessment of local disease extent; accuracy is limited and is generally restricted to determining whether the tumor invades the seminal vesicles. It is also used to determine the size of the gland for the calculation of PSAD and to guide the placement of radioactive seeds during implantation.

Pathology The noninvasive proliferation of epithelial cells within ducts is termed *prostatic intraepithelial neoplasia* (PIN). It is considered the precursor of cancer, but not all PIN lesions develop into invasive cancers. Nevertheless, on a genetic level, these regions are highly unstable and typically multifocal. Of the cancers identified, >95% are adenocarcinomas; the remainder include squamous cell tumors, transitional cell tumors, and, rarely, carcinosarcomas. Metastases to the prostate are rare, but in some cases, transitional cell tumors originating in the bladder or colonic lesions may invade the gland directly. In the evaluation for adenocarcinoma, each core is examined for the presence or absence of cancer. When cancer is identified, the extent and grade are assessed and the presence or absence of perineural invasion or extracapsular extension reported. Histologic grade is based most commonly on the *Gleason system*, in which the dominant and secondary glandular histologic patterns are independently assigned numbers from 1 to 5 (best to least differentiated) and summed to give a total score of 2 to 10 for each tumor. The grading is reproducible and correlates with clinical outcomes. The most poorly differentiated area of tumor (i.e., the area with the highest histologic grade) often determines biologic behavior.

STAGING

The TNM staging system (Table 95-2) includes categories for cancers identified solely on the basis of an abnormalPSA with no palpable abnormalities onDRE(T1c), those that are palpable but clinically confined to the gland (T2), and those that have extended outside of the gland (T3 and T4). The presence or absence of nodal (N) and distant metastases (M) are also recorded. Clinical staging alone is inaccurate in assessing capsular invasion and the probability of spread to nodal or more distant sites. To refine this assessment, the TNM system has been modified to incorporate the results of imaging studies such as ultrasound or magnetic resonance imaging (MRI) in the assignment of T stage.

Computed tomography (CT) scans lack sensitivity and specificity to detect extraprostatic extension and in visualization of lymph nodes. MRI is an improvement, particularly with an endorectal coil and is superior to CT. T1-weighted images demonstrate the periprostatic fat, periprostatic venous plexus, perivesicular tissues, lymph nodes, and bone marrow. T2-weighted images demonstrate the internal architecture of the prostate and seminal vesicles. Most cancers have a low signal, while the normal peripheral zone has a high signal. Nevertheless, MRI lacks sensitivity and specificity. No single test accurately predicts pathologic stage at surgery.

Another limitation of the TNM system is that the majority of men are now being diagnosed with T1c or T2 disease. Thus, to refine the prediction of local disease extent, most groups are now using multiplex staging models based on a combination of the findings of theDRE, biopsy, Gleason score, and baselinePSA(Table 95-3). Others are developing models based on the number of cores and the percentage of each core

involved by tumor. This information can be used to assist patients in selecting treatments, although it remains controversial how recommendations should be affected by a particular level of probability of node-positive disease.

These same parameters are also being used to assess the probability of cure (Fig. 95-2). Some tumors that have extended beyond the confines of the gland may still be curable, while others that are still organ-confined may not. Because successful surgery removes all prostatic tissues, both benign and malignant, and radiation therapy eliminates only the malignant component, different definitions of cure are needed depending on the modality used. Thus, following surgery, thePSAlevel should become undetectable; following radiation therapy, it should generally fall to <1.0 ng/mL. What the models do not address is what probability of cure a patient would accept to proceed with a given approach. For example, would one categorically deny a surgical procedure to a 40-year-old male if the probability of cure was only 15%? These same models can also be used to stratify patients into risk groups to assess outcomes of specific therapies. While this form of analysis does not replace prospective trials, it does eliminate the bias associated with the tendency to refer older and more infirm patients for radiation therapy and younger, healthier individuals for surgery.

To complete the staging evaluation, patients may undergo radionuclide bone scanning. This test is highly sensitive but relatively nonspecific, because areas of increased uptake are not always secondary to osteoblastic activity from metastases. Healing fractures, arthritis, Paget's disease, and numerous other conditions will also show abnormal uptake. True-positive bone scan results are rare if thePSA is<8 and uncommon when the PSA is<10 ng/mL. More common is a false-positive scan, which, in turn, leads to additional low-yield testing.CT scans yield little useful clinical information unless the probability of lymph node metastases is >30% using nomogram predictions; anMRI is more likely to detect pathologically significant nodal disease. Molecular diagnostics are being performed that seek to identify the presence of circulating prostate cancer cells using an assay for PSA based on reverse transcriptase polymerase chain reaction (RT-PCR) in the leukocyte fraction of the peripheral blood or bone marrow. A large proportion of men with tumors seemingly confined to the organ test positive; the significance is unclear. These procedures are in their infancy, and application is not advised on a routine basis.

TREATMENT SELECTION: THE MODEL OF CLINICAL STATES

The framework for evaluating the risks from an enlarging but nonmalignant gland, the probability that a clinically significant cancer is present in an individual with or without urinary symptoms, and the probability that a patient with cancer will develop symptoms or die of prostate cancer are provided by the clinical states model illustrated in Fig. 95-3. It includes clinically significant milestones where interventions might be considered and allows for the assessment of the prognosis of the treated patient. In the first state are patients with no cancer diagnosis. It includes patients with benign proliferative disorders or those who warrant screening on the basis of family history or a level ofPSA or symptoms. In the second state are those with a cancer that is clinically confined to the gland. For these patients the issue is to determine which tumors require treatment based on their biologic potential, which can be eradicated by local means alone, and which require a combined-modality approach that includes systemic therapy to effect

cure. The third state includes those who have a rising PSA level after surgery or radiation for localized disease but who have no clinically detectable lesions on scans. Next are patients with detectable metastases who have not undergone castration, and the last level is those who have detectable disease on scan despite castration. The risk of death from cancer relative to the risk of death from comorbid conditions increases over time, being greatest for the patient who has progressed after hormonal therapy.

At any point, a patient resides in only one state and remains there until the disease progresses. Thus, a patient who presents with a localized prostate cancer who has had all cancer removed surgically remains in the state of localized disease as long as hisPSAremains undetectable. In this way, both time factors and that the fact that a patient has been treated are accounted for. Overall treatment effects are assessed by measuring time within a particular state. The scheme also allows a distinction between cure, elimination of all cancer cells, with an undetectable PSA and cancer control, i.e., modulating the rate of growth so that the patient dies of other causes. In this paradigm, a patient with a detectable PSA who dies of other causes having suffered no morbidity from the disease or its treatment, PSA level, is considered a therapeutic success.

MANAGEMENT BY STATES

NO CANCER DIAGNOSIS

Screening The American Cancer Society (ACS) and the AUA recommend an annualDRE and a determination ofPSAlevel for all men aged 50 to 79. Individuals with a first-degree relative with prostate cancer and African Americans, who have a higher risk of dying of the disease, are advised to begin testing at age 45. Routine screening for prostate cancer has been advised despite a lack of prospective, randomized, controlled trials proving the benefit of the approach because the disease rarely causes symptoms until it is advanced. The more widespread use of routine DREs and PSA testing has resulted in a significant increase in the proportion of men with clinically localized tumors, a reduced frequency of nodal spread, and a decreased frequency of nodal and osseous disease at presentation. Risks of screening are unnecessary morbidity or mortality from overdetection and overtreatment. Formal clinical trials are underway, but until the studies are complete and the results available, men must make an informed decision to be screened or not.

Hyperplasia A patient with an enlarged prostate who has no symptoms and normalPSAlevels generally does not require treatment. Those with symptoms such as an inability to urinate, renal insufficiency, urinary tract infection, gross hematuria, or bladder stones are candidates for prostate surgery. As the natural history is not well defined, it is not always clear whether to intervene and, if so, how. The majority of men do not develop significant obstruction, and in many, minor irritative and/or obstructive symptoms change slowly or not at all. In these cases urine flow studies can identify those whose maximum flows are normal and who are unlikely to benefit from treatment. Measuring postvoid residual volume identifies patients likely to fail a "watch and wait" approach, while pressure-flow studies may identify those with primary bladder dysfunction. A cystoscopic examination is advised for all patients with hematuria and to assess the urinary outflow tract before a surgical intervention. Imaging of the upper urinary tract by ultrasonography or intravenous pyelography should be reserved for

patients with indications such as hematuria, a history of stones, or prior urinary tract problems.

Most patients are monitored and/or treated medically, after a discussion with their physicians about the degree of incapacity and/or discomfort present and the likely outcome of each potential treatment strategy. A variety of decision diagrams have been proposed. Patients who opt for deferred therapy should be evaluated on an annual basis by the reassessment of symptoms and clinical manifestations. Medical therapies include finasteride, which blocks the conversion of testosterone to dihydrotestosterone, the principal androgen in the prostate, by competitively inhibiting the 5a-reductase enzyme. A dose of 5 mg/d causes an average decrease in prostate size of ~24%, an increase in urine flow rates, and, in some, improvement in symptoms. Long-term efficacy has not been documented, but symptomatic improvement has been documented for[3]3 years provided therapy is continued.a-Adrenergic blockers such as terazosin act by relaxing the smooth muscle of the bladder neck, increasing peak urinary flow rates and reducing symptoms. No data prove that these agents influence the progression of the disease.

Patients who do not improve or who progress on medical therapy require surgical intervention. Surgical approaches include a transurethral resection of the prostate (TURP); transurethral incision; or removal of the gland by a retropubic, suprapubic, or perineal route. Other approaches include ultrasound, coils, stents, lasers, or hyperthermia. Overall, surgery offers the best chance for improving symptoms, at the cost of the highest rate of complications. TURP is the most common surgical procedure. Transurethral incision of the prostate is of similar efficacy in men with relatively small prostates and can be performed in ambulatory settings. Open prostatectomy is usually reserved for men with massive prostates; it has the longest recovery time and the highest morbidity, particularly impotence. Transurethral incision has the least morbidity overall and is least disruptive to ejaculatory function.

ElevatedPSAand No Cancer Diagnosis on Biopsy Patients who have undergone a biopsy procedure and do not have a cancer diagnosis should continue to be monitored. In some cases, a repeat biopsy session with particular attention to the transition zone is advised. The frequency ofPIN is similar in men of different ethnic backgrounds around the world, while the incidence of the clinical disease varies in different ethnic groups. Prevention of progression from PIN to cancer is an area of active research. Proving the benefit of a prevention strategy is difficult because of the long-term follow-up that is necessary, the large sample sizes required to demonstrate a difference in outcome, and the absence of surrogate measures that predict for efficacy. Agents under study include the retinoids, vitamin D, selenium, soy, and modifications of dietary fat. Most are based on epidemiologic data suggesting a decreased prostate cancer risk. A large-scale, double-blind, randomized, multicenter trial of finasteride in men over age 55 has accrued 18,000 men, and follow-up is awaited.

CLINICALLY LOCALIZED DISEASE

Localized prostate cancer (stages T1-2, NX or 0, M0) may require no therapy, may be curable with localized therapy, or may require combined-modality systemic and local therapy. The key is to distinguish these distinct prognostic groups. Treatment planning

includes an assessment of the probability of local control, local failure, and systemic failure. The more advanced the disease, the lower the probability of local control and the higher the probability of systemic relapse. In general, these tumors are managed by watchful waiting, radical surgery, or radiation therapy. Comparisons between these approaches are limited by the lack of prospective comparative trials, referral biases, and differences in the endpoints evaluated.

Conservative Management (Watchful Waiting) The concept of watchful waiting, or deferred therapy, evolved from the recognition of the high prevalence of the disease in the population, the low probability that some cancers would affect an individual's quality-adjusted life expectancy, and the fact that morbidities associated with the local treatment options were unacceptable to many patients. Watchful waiting acknowledges the facts that the natural history of an untreated prostate cancer is to progress and that it may be difficult to monitor progression within the gland so that the "window of curability" is not lost. That the disease is often multifocal leads to the possibility that the biopsy on which the decision to defer therapy is made may not represent accurately the malignant potential of a second unidentified cancer. Within 10 years of diagnosis, most tumors produce local symptoms such as urinary retention, incontinence, hematuria, ureteral and bowel obstruction, and pelvic pain, but these complications rarely lead to the death of the patient. Some tumors may metastasize, but few patients succumb to the disease. Case selection criteria are evolving, but in general, watchful waiting is not advised for patients with high-grade disease or for those with a>10-year life expectancy. Some physicians consider observation only for patients with low-grade tumors (Gleason score£6) that do not involve more than a small percentage of a single core.

Radical Prostatectomy The objective of a radical prostatectomy is the removal of all prostate tissue with a clear margin of resection, preservation of the external sphincter to maintain continence, and sparing of the autonomic nerves in the neurovascular bundle so that potency is retained. The procedure is performed through a retropubic or perineal approach. In contemporary series, hospital stays are short; mortality <0.4%; and complications such as rectal injury, deep vein thrombosis, and embolic events are rare. The procedure is recommended primarily for patients with clinically localized disease (T1c-T3a, N0 or NX, M0 or MX) who have a life expectancy of>10 years. The operation is not justified in men with a life expectancy <5 years. Properly performed, the procedure requires appropriate case selection and meticulous technique that permits the delineation of the anatomy of the gland and surrounding tissues. In one review, the overall rate of positive margins was 25%. Careful planning can reduce this rate. For example, by considering the laterality and extent of disease, it may be apparent that nerve-sparing cannot be achieved without compromising cancer control. A positive margin increases the risk of progression significantly. ThroughPSA-based detection, the proportion of men with positive nodes and positive margins continues to decline.

Complication rates, specifically the probability of developing a bladder neck contracture, incontinence, or impotence, vary depending on the experience of the surgeon and whether the patient or the physican is describing the outcome. Rates of incontinence based on physician reporting are 5 to 10%, compared to 19 to 31% based on independent questioning by a third party. Time is also a consideration, as full recovery of function may not occur for weeks or months following the procedure. Factors associated with incontinence include older age, functional length of the urethra, surgical

technique, preservation of neurovascular bundles, and development of an anastomotic stricture. If the nerves are preserved, ~70% of men recover the ability to achieve an erection sufficient for penetration. Most men are impotent immediately after the procedure and gradually recover function over 6 to 12 months. Often the quality of the erection is decreased from preoperative levels. Nevertheless, with orally active drugs such as sildenafil, intraurethral inserts of alprostadil, and intracavernosal injections of vasodilators, many patients can achieve nearly natural erections and recover satisfactory sexual activity. Factors associated with recovery include younger age, quality of erections before the operation, and the absence of damage to the neurovascular bundles. Loss of one bundle is associated with a 75% reduction in the recovery of function.

After a successful radical prostatectomy in which all prostate tissue has been removed, serumPSAlevels should become undetectable within 4 weeks, based on the half-life of 3 days. If the PSA level remains detectable or becomes detectable after having been undetectable, the patient is considered to have persistent disease or to have a recurrence. In the absence of adjuvant treatment, most patients destined to recur do so within the first 5 years after surgery. Thus, one early benchmark of "success" is the probability of freedom from PSA (or "biochemical") progression. This varies as a function of initial clinical stage, Gleason grade, and serum PSA level before surgery. In one series of 1359 men with clinical stages T1/T2 cancer followed for a mean of 44 months (range 1 to 170), the PSA relapse-free survival rates were 78% at 5 years and 73% at 10 years. In a separate series of T1c patients, 89% were free of progression at 5 years. Considered by baseline PSA levels, 95% of those with a normal level (<4 ng/mL) and 68% of those with a PSA>10 ng/mL were free of progression at 5 years. Considered by grade or Gleason sum in the biopsy specimen, 5- and 10-year PSA relapse-free survivals were 56% and 46%, respectively, for those with tumors of Gleason score of 7, and 46 to 53% at 5 years for those with tumors of Gleason score ³8. These outcomes appear superior to those reported with watchful waiting, recognizing the limitations in comparing the results of nonrandomized selected series.

The most significant predictor of recurrence is pathologic stage. When the disease is confined to the organ and has not extended into the periprostatic soft tissue, 91 to 97% of patients remain free of progression at 5 years and 85 to 92% at 10 years. Extension to the periprostatic soft tissues (pT3a and N0) decreasesPSArelapse-free probabilities to 74% and 68% at 5 and 10 years, respectively, which is decreased further to 40 to 47% and 25% if there is seminal vesicle invasion (pT3c and N0). The high frequency of extracapsular extension and positive surgical margins in patients with clinically localized prostate cancers that were presumed to be confined to the gland led to the investigation of neoadjuvant hormonal therapy. The results of several large contemporary series evaluating 3 months of hormone therapy before surgery showed that, on average, positive margins are reduced from 41% to 17%, serum PSA levels by 96%, and prostate volume by 34% with neoadjuvant hormone therapy. The surrogate of a reduction in positive margin rates was not predictive of a reduction in failure rates, as the time to PSA relapse was no different between groups receiving and not receiving hormones. As such, neoadjuvant hormonal therapy is not recommended. Several recurrence models are available that incorporate all of these factors.

Radiation Therapy Radiation therapy can be delivered externally, by implantation of

radioactive sources into the gland, or a combination of both. As is the case with surgery, outcomes vary as a function of the method, the dose, the endpoints, and whether outcomes were based on clinical or pathologic staging of the lymph nodes. Some groups report local control, and othersPSArelapse-free survival, time to metastases, or overall survival. Cause-specific survivals are rarely reported. Local control can be reported on the basis of aDREalone or the more stringent criterion of a negative biopsy at 18 to 24 months following treatment. Length of follow-up can also influence the results. To standardize reporting, the American Society of Therapeutic Radiation Oncology has developed a consensus definition of PSA relapse as three consecutive rising PSA values from the nadir value.

External Beam Therapy Overall, outcomes with external beam therapy are similar for patients with T1 and T2a disease to those obtained with radical surgery. Outcomes for patients with locally advanced disease (T2bc and T3/T4) are less favorable, the result of both inadequate control of the primary tumor and the high rate of systemic failure associated with more advanced disease. For the latter group, 30 to 40% of patients relapse locally using standard doses.

Conventional techniques use simulators and CTscans of the pelvis to determine the location and shape of the target volume and the surrounding normal organs. Typical treatment plans use a four-field pelvic box designed to include the prostate, seminal vesicles, and the locally draining lymph nodes. Normal structures are protected by shaping the beams with cerrobend trim blocks. Therapy is delivered on a daily basis, excepting weekends, in 1.8- to 2.0-Gy fractions. Outcomes are dose-dependent; in one series of stage C patients, actuarial 7-year local recurrence rates were 36% for those receiving 60 to 64.9 Gy, 32% for those receiving 65 to 69.9 Gy, and 24% for those treated at ³70 Gy. Complication rates also increase with increasing dose. Using standard doses, grade 2 or greater rectal and/or urinary symptoms requiring medication occur in 60% of cases, while late sequelae such as cystitis, hematuria, stricture, or bladder contracture occur in 7% of cases. The frequency of adverse events is significantly higher in patients who have undergone aTURP, while the frequency of rectal complications is directly related to the volume of the anterior rectal wall receiving full-dose treatment. The frequency of erectile dysfunction is related to the quality of erections before treatment, the dose administered, and the time of assessment. Impotence is related to a disruption of the vascular supply and not the nerve fibers.

More contemporary approaches use three-dimensional conformal radiation therapy (3D-CRT) techniques with sophisticated computer-generated treatment plans to deliver the prescribed radiation dose to the entire target volume, while conforming to the anatomic boundaries of the tumor in its entire three-dimensional configuration. This treatment method has increased the ability to control the cancer through the administration of higher radiation doses, with less morbidity to the surrounding normal organs. In a series of 743 patients treated with 3D-CRT, 90% of patients receiving 75.6 or 81.0 Gy achieved aPSAnadir of £1 ng/mL compared with 76% and 56% of those treated with 70.2 Gy and 64.8 Gy, respectively ($p < .001$).

As is the case with surgically treated patients, pre-therapy nomograms can be used to stratify patient groups. In one series, the 5-year actuarialPSArelapse-free survival for patients with favorable prognostic indicators (stage T1/T2, pretreatment PSA of 10.0

ng/mL, and Gleason score of 6) was 85%; it was 65% for those with an intermediate prognosis (one of the prognostic indicators with a higher value) and 35% for those with unfavorable features (two or more indicators with higher values) (p <.001).

Tolerance of3D-CRThas been excellent despite the use of higher radiation doses; grade 3 to 4 rectal or urinary toxicities were seen in 2.1% of patients. In contrast, among patients treated with conventional external-beam radiotherapy, the incidence of grade 3 to 4 toxicities for patients who received radiation doses of>70 Gy was 6.9%.

To improve outcomes for patients with unfavorable features, several groups have explored hormone therapy before radiation therapy. Prospective randomized trials showed improved local control and a delay in time toPSArelapse in patients receiving 2 to 3 years of treatment. The impact on survival has been less clear.

Interstitial Therapy Interstitial brachytherapy is based on the principle that the deposition of radiation energy in tissues decreases exponentially as a function of distance from the radiation source. By infiltrating tumor tissue with radioactive sources, intensive irradiation is delivered to the prostate with minimal irradiation of the surrounding tissues. In a series of 197 patients followed for a median of three years, 5-year actuarialPSArelapse-free survival for patients with pre-therapy PSA levels of 0 to 4, 4 to 10, and >10 were 98%, 90%, and 89%, respectively. Nevertheless, many physicians feel that implantation is best reserved for patients with good or intermediate prognostic features.

Overall, the procedure is well tolerated, although most patients experience urinary frequency and urgency, which can persist for several months. Incontinence has been seen in 2 to 4% of cases. Higher complication rates are observed in patients who have undergone a priorTURP or who have obstructive symptoms at baseline. Proctitis has been reported in<2% of patients. Longer follow-up will be necessary to see whether the overall frequency of impotence is lower, higher, or the same as that observed using external radiation delivery techniques.

RISING PSA

Included in the group of patients with a risingPSA and no evidence of metasatic disease on scans are those who have progressed after watchful waiting, radical prostatectomy, radiation therapy, or both surgery and radiation, with or without prior hormone exposure. For these individuals, the issue is to determine whether the rising PSA is due to local persistence or recurrence (additional therapy to the primary site might be curative) or the result of micrometastatic disease. Imaging studies such asCT,MRI, or bone scan are typically uninformative. The objective is to assess the probability of disease progressing to the point where metastases will occur or cause symptoms. This is the point in the disease where the probability of death from disease exceeds the probability of death from other causes. Difficulty in making these predictions comes from the fact that most patients with a rising PSA receive some form of therapy before the development of metastatic disease, making it virtually impossible to assess the natural history.

To estimate the probability of having a local or systemic recurrence, many investigators use the time toPSAfailure or the rate of rise of PSA as predictive factors. In general,

recurrences documented >1 year after primary treatment tend to be localized, while those recurring in<1 year tend to be systemic. These predictions are not hard and fast. In one series of patients with PSA recurrence after surgery who did not receive systemic therapy until metastatic disease was documented, the median time to metastatic progression was 8 years, and 63% of the patients with rising PSA values remained free of metastases at 5 years. Patients with tumors of Gleason score 8 to 10 had a probability of metastatic progression of 37%, 51%, and 71% at 3, 5, and 7 years, respectively. Combining a high-grade histology and rapid PSA doubling time, the proportion with metastases was 23%, 32%, and 53% during the same time intervals if the time to recurrence was<2 years and the PSA doubling time was>10 months and 47%, 69%, and 79% for the same recurrence interval with <10 months doubling time. For those with tumors of Gleason score 5 to 7, a PSA recurrence in the first 2 years and a doubling time of <10 months identified a group of patients with a frequency of metastases of 19%, 65%, and 85% at 3, 5, and 7 years, respectively.

Prostascint scanning uses a radioactive antibody to prostate-specific membrane antigen (PSMA), which is highly expressed on prostate epithelial cells. For a patient who has undergone a radical prostatectomy, antibody localization to the prostatic fossa is suggestive of local recurrence, in which case external beam radiation therapy might be recommended. Others recommend that a biopsy of the urethrovesical anastamosis be obtained before considering radiation. Most, however, rely on clinical criteria with the additional caveat that the probability of durablePSAcontrol varies inversely with the level of PSA at the start of radiation therapy. Radiation therapy is usually not recommended if the PSA level exceeds 1 to 2 ng/mL or if the PSA was persistently elevated after surgery (indicating that disease-free status was not achieved). For patients with a rising PSA after radiation therapy, a salvage prostatectomy can be considered if (1) residual disease is detected in the gland based on a repeat biopsy, (2) the tumor was amenable to surgical extirpation before radiation therapy, and (3) metastatic disease is absent on imaging studies. Unfortunately, case selection is poorly defined in most series, and morbidities have been significant. As currently performed, virtually all patients are impotent, and ~45% have either incontinence or stress incontinence. Bleeding, bladder neck contractures, and rectal injury are not uncommon.

METASTATIC DISEASE

Noncastrate The removal or blockade of androgens by medical or surgical means is the mainstay of treatment for patients with advanced disease. Surgical orchiectomy is the "gold standard" but is the least preferred by patients. Medical therapies can be subdivided into those that result in a lowering of serum testosterone levels, e.g., gonadotropin-releasing hormone (GnRH) agonists and antagonists and estrogens, and the antiandrogens (Fig. 95-4). Inhibitors of adrenal enzyme synthesis such as ketoconazole and aminoglutethimide are typically used as second-line treatment. The antitumor effects of agents that lower serum testosterone levels are similar, but toxicities differ. Castration is associated with gynecomastia, impotence, weakness, fatigue, hot flushes, loss of muscle mass, changes in personality, anemia, depression, and loss of skeletal mass. Loss of bone mass can be reduced by coadministration of bisphosphonates.

GnRHanalogues (leuprolide acetate and goserelin acetate) initially produce a rise in

luteinizing hormone (LH) and follicle-stimulating hormone (FSH), followed by a downregulation of receptors in the pituitary gland, which effects a chemical castration. The initial rise in testosterone may result in a clinical flare of the disease. As such, these agents are contraindicated in men with significant obstructive symptoms, cancer-related pain, or spinal cord compromise. The flare can be prevented by pretreatment with antiandrogens. Pure GnRH antagonists that do not produce the initial rise in testosterone will shortly be available.

Diethylstilbestrol (3 mg/d) produces castrate levels of testosterone in 1 to 2 weeks and is inexpensive. Its significant cardiovascular toxicities include edema, congestive heart failure, myocardial infarction, cerebrovascular accidents, phlebitis, and pulmonary embolism. Gynecomastia, a common adverse event, can be reduced by prophylactic irradiation of the breasts. Progestational agents such as medroxyprogesterone acetate (Provera) and megestrol acetate (Megace), are inferior to conventional castration and are not used as first-line treatment. The antifungal agent ketoconazole administered at a dose of 1200 mg/d (six times the antifungal dose) produces a chemical castration in 24 h. It is absorbed in an acid environment and typically prescribed with citrus juices to improve absorption; antacids or H_2blocking agents reduce absorption and should be avoided when the pills are administered. The effects on testosterone synthesis, however, are not sustained, and long-term use is limited by hepatoxicity. It can be useful for the unusual patient who presents with a coagulopathy or spinal neurologic compromise and who requires a rapid response. Aminoglutethimide, a second adrenal synthesis inhibitor, was originally developed as an antiseizure medication. It is administered with hydrocortisone. Side effects include somnolence, fatigue, rash, and, after prolonged periods, hypothyroidism; its use is limited.

Nonsteroidal antiandrogens such as flutamide (Eulexin), bicalutamide (Casodex), or nilutamide (Anandron) block the binding of androgens to the receptor. They do not block the production ofLHcentrally, and as a result, serum testosterone levels increase. These drugs have been used clinically in several situations: (1) to block the flare from the initial rise in testosterone fromGnRHuse, (2) as monotherapy to preserve potency, and (3) as part of a combined androgen-blockade approach designed to simultaneously inhibit testicular and adrenal androgens. Toxicities differ among the agents but generally include gynecomastia (which can be significant), fatigue, elevations in serum transaminases, and diarrhea. The latter is the most common reason the drug is discontinued. Nilutamide is also associated with impaired adaptation to darkness, alcohol intolerance, and, rarely, pneumonitis.

Hormonal therapy is the treatment of choice, but the timing of treatment is not as clear. Early administration of hormones delays progression, but the survival benefit is less clear. It is controversial whether hormonal therapy should be initiated with a risingPSAlevel or whether treatment should be held until metastatic disease is detectable on scans. Most physicians recommend use of hormonal therapy with PSA elevation, based on evidence from clinical trials suggesting a survival advantage to such an approach.

A second controversy is whether a combined androgen-blockade that includes an antiandrogen is superior to castration without an antiandrogen. In randomized comparisons, both positive and negative trials have been reported but the majority have

shown no difference. Some feel that the method of primary castration (GnRHanalogue vs. orchiectomy) and the class of antiandrogen (steroidal vs. nonsteroidal) may influence outcome, but this is controversial. A meta-analysis of 22 randomized trials that included 5710 patients showed an absolute 2.1% difference in mortality at 5 years. This translated into a 6.4% reduction in the annual odds of death. A similar analysis by the Blue Cross/Blue Shield Association Evidence-Based Practice Center concluded that no benefit was seen with combined androgen blockade.

Castrate The management of patients who progress on hormone therapy requires documentation of a castrate status and an evaluation for residual hormone sensitivity. For patients on hormonal therapy, this involves discontinuing all hormonal therapy to evaluate for a withdrawal response. Responses are noted within a few weeks of stopping the medication, with the exception of nilutamide and bicalutamide, both of which have a long terminal half-life. For these agents, the response may be delayed. Signs of benefit include declines inPSAlevel, regression of measurable disease, palliation of pain, and improvements in cancer-related anemia.

Patients who are documented to be castrated, and/or who progress after a trial of withdrawal, are often given one additional hormonal manipulation. Depending on prior hormone exposure, options include inhibitors of adrenal steroid hormone synthesis such as ketoconazole and aminoglutethimide, glucocorticoids, antiandrogens, estrogens, or progestational agents. The responses are often short-lived and do not occur in the majority of patients. Nevertheless, the response can be durable in some patients and provide significant palliation in the absence of curative therapies. Glucocorticoids have been associated with clinical benefit and declines inPSAlevels in 30 to 40% of cases.

Estramustine (Emcyt) is a synthetic combination of estrogen with a nitrogen mustard moiety at C17 that affects microtubule assembly and disassembly. It has no alkylating effects in vivo. About 20% of the drug is metabolized to pure estrogenic moieties, which exert an antigonadotropin effect and which account for the side effect profile. It is often used in patients who have failed hormone therapy and has additive/synergistic effects with other drugs including the vinca alkaloids (vinblastine and navelbine), taxanes (paclitaxel, docetaxel), and the podophyllotoxins (etoposide).

Patients who progress on primary hormone therapy and after hormone withdrawal and receive one additional (second-line) intervention are considered to have "hormone-independent" or "hormone-refractory" disease. At this point, chemotherapy is often considered, although some feel it has no role in the management of prostate cancer because no single agent or combination of drugs has been shown to improve survival in a prospective randomized trial. In the absence of a proven benefit in survival, it is important to consider the specific goals of therapy before it is recommended. These goals might include palliation of symptoms, delaying progression, or inducing decreases in thePSAlevel. Interpreting reported outcomes with individual agents is limited in part by differences in case selection and the wide range of endpoints used.

Mitoxantrone has modest activity as a single agent and is more effective at relieving pain and improving quality of life when given together with prednisone. No effect on survival has been shown. However, mitoxantrone plus prednisone and estramustine plus vinblastine are often used to palliate symptoms of disease. The most frequently

utilized contemporary regimens are weekly combinations of estramustine and a taxane (paclitaxel or docetaxel). Weekly doxorubin also provides palliation.

PALLIATION OF PAIN

Pain is one of the most feared debilitating manifestations of advanced disease. Palliation of pain can be achieved with external beam radiation therapy, bone-seeking radioisotopes, cold bisphosphonates, and chemotherapy. The goals are to relieve symptoms, prevent complications, and improve quality-adjusted life expectancy. To optimize treatment selection, it is important to consider the sites and distribution of the pain and the presence or absence of neurologic compromise. Spinal cord compression is one of the most devastating complications. Once a loss of function is documented, the probability of recovery is small. Particular areas where a high index of suspicion is required are the base of the skull, which can produce a variety of symptoms including diplopia, deafness, difficulty swallowing, dysarthria, and facial weakness; and mental nerve compression in the jaw, resulting in a numb lip and chin, which can interfere with eating. In these situations, external beam radiation together with glucocorticoids are required.

For a solitary lesion that is symptomatic and can be treated through a single port, external beam radiation therapy is the treatment of choice. Depending on the clinical situation, a single high-dose fraction (6 to 9 Gy) may be all that is necessary. The overall utility is limited by the facts that metastases are rarely solitary and that additional untreated areas often become symptomatic in a relatively short time. In other cases, wide-field portals are needed. For those with more diffuse disease, bone-seeking radioisotopes are available.^{153}Sm-EDTMP (quadramet) emits b particles and a photon; its half-life is about 22 h. When given at 37 MBq/kg (1 mCi/kg), half the administered dose goes to bone, and 70 to 95% of patients experience decreased bone pain within 2 weeks that lasts 8 to 15 weeks.^{89}Sr (metastron) is a pure b-emitter with a half-life of 50 days that emits a 1.46-MeV electron with a 2- to 3-mm range in bone. Used alone it has modest effects on the level ofPSA but does provide a degree of palliation that is similar to external beam approaches. The results of randomized comparisons suggest a systemic effect, as fewer patients treated with the isotope developed new areas of pain or required additional radiation therapy compared to patients receiving radiation therapy alone. Cold bisphosphonates have been shown to be superior to placebo in the prevention of skeletal events for patients with breast cancer, lung cancer, and multiple myeloma. They may be active in prostate cancer as well.

ACKNOWLEDGEMENT

Dr. Jean Wilson and Dr. Arthur Sagalowski were the authors of this chapter in the 14th edition and parts of their chapter have been retained here.

(Bibliography omitted in Palm version)

Back to Table of Contents

96. TESTICULAR CANCER - *Robert J. Motzer*, *George J. Bosl*

Primary germ cell tumors (GCTs) of the testis, arising by the malignant transformation of primordial germ cells, constitute 95% of all testicular neoplasms. Infrequently, GCTs arise from an extragonadal site, including the mediastinum, retroperitoneum and, very rarely, the pineal gland. This disease is notable for the young age of the afflicted patients, the totipotent capacity for differentiation of the tumor cells, and its curability;>90% of all newly diagnosed patients will be cured. Experience in the management of GCTs leads to improved outcome.

INCIDENCE AND EPIDEMIOLOGY

Nearly 6900 new cases of testicularGCT were diagnosed in the United States in 2000; the incidence of this malignancy has increased slowly over the past 40 years. The tumor occurs most frequently in men between the ages of 20 and 40. A testicular mass in a man 50 years or older should be regarded as a lymphoma until proved otherwise. GCT is at least 4 to 5 times more common in white than in African-American males, and a higher incidence has been observed in Scandinavia and New Zealand than in the United States.

ETIOLOGY AND GENETICS

Cryptorchidism is associated with a severalfold higher risk ofGCT. Abdominal cryptorchid testes are at a higher risk than inguinal cryptorchid testes. Orchiopexy should be performed before puberty, if possible. Early orchiopexy reduces the risk of GCT and improves the ability to save the testis. An abdominal cryptorchid testis that cannot be brought into the scrotum should be removed. About 2% of men with GCTs of one testis will develop a primary tumor in the other testis. Testicular feminization syndromes increase the risk of testicular GCT, and Klinefelter's syndrome is associated with mediastinal GCT.

An isochromosome of the short arm of chromosome 12 [i(12p)] is pathognomonic forGCT of all histologic types. Excess 12p copy number either in the form of i(12p) or as increased 12p on aberrantly banded marker chromosomes occurs in nearly all GCT, but the gene(s) on 12p involved in the pathogenesis are not yet defined.

CLINICAL PRESENTATION

A painless testicular mass is pathognomonic for a testicular malignancy. More commonly, patients present with testicular discomfort or swelling suggestive of epididymitis and/or orchitis. In this circumstance, a trial of antibiotics is reasonable. However, if symptoms persist or a residual abnormality remains, then testicular ultrasound examination is indicated.

Ultrasound of the testis is indicated whenever a testicular malignancy is considered and for persistent or painful testicular swelling. If a testicular mass is detected, a radical inguinal orchiectomy should be performed. Because the testis develops from the gonadal ridge, its blood supply and lymphatic drainage originate in the abdomen and descend with the testis into the scrotum. An inguinal approach is taken to avoid

breaching anatomic barriers and permitting additional pathways of spread.

Back pain from retroperitoneal metastases is common and must be distinguished from musculoskeletal pain. Dyspnea from pulmonary metastases occurs infrequently. Patients with increased serum levels of human chorionic gonadotropin (hCG) may present with gynecomastia. A delay in diagnosis is associated with a more advanced stage and possibly worse survival.

The staging evaluation for GCTincludes a determination of serum levels of afetoprotein (AFP) andhCG. After orchiectomy, a chest radiograph and a computed tomography (CT) scan of the abdomen and pelvis should be performed. A chest CT scan is required if pulmonary nodules, mediastinal or hilar disease is suspected. *Stage I disease* is limited to the testis, epididymis, or spermatic cord. *Stage II disease* is limited to retroperitoneal (regional) lymph nodes. *Stage III disease* is disease outside the retroperitoneum, involving supradiaphragmatic nodal sites or viscera. The staging may be "clinical" -- defined solely by physical examination, blood marker evaluation, and radiographs -- or "pathologic" -- defined by an operative procedure.

The regional draining lymph nodes for the testis are in the retroperitoneum, and the vascular supply originates from the great vessels (for the right testis) or the renal vessels (for the left testis). As a result, the lymph nodes that are involved first by a right testicular tumor are the interaortocaval lymph nodes just below the renal vessels. For a left testicular tumor, the first involved lymph nodes are lateral to the aorta (para-aortic) and below the left renal vessels. In both cases, further nodal spread is inferior and contralateral and, less commonly, above the renal hilum. Lymphatic involvement can extend cephalad to the retrocrural, posterior mediastinal, and supraclavicular lymph nodes. Treatment is determined by tumor histology (seminoma versus nonseminoma) and clinical stage (Table 96-1).

PATHOLOGY

GCTsare divided into nonseminoma and seminoma subtypes. Nonseminomatous GCTs are most frequent in the third decade of life and can display the full spectrum of embryonic and adult cellular differentiation. This entity comprises four histologies: embryonal carcinoma, teratoma, choriocarcinoma, and endodermal sinus (yolk sac) tumor. Choriocarcinoma, consisting of both cytotrophoblasts and syncytiophoblasts, represents malignant trophoblastic differentiation and is invariably associated with secretion ofhCG. Endodermal sinus tumor is the malignant counterpart of the fetal yolk sac and is associated with secretion ofAFP. Pure embryonal carcinoma may secrete AFP or hCG, or both; this pattern is biochemical evidence of differentiation. Teratoma is composed of somatic cell types derived from two or more germ layers (ectoderm, mesoderm, or endoderm). Each of these histologies may be present alone or in combination with others. Nonseminomatous GCTs tend to metastasize early to sites such as the retroperitoneal lymph nodes and lung parenchyma. One-third of patients present with disease limited to the testis (stage I), one-third with retroperitoneal metastases (stage II), and one-third with more extensive supradiaphragmatic nodal or visceral metastases (stage III).

Seminoma represents about 50% of allGCTs, has a median age in the fourth decade,

and generally follows a more indolent clinical course. Most patients (70%) present with stage I disease, about 20% with stage II disease, and 10% with stage III disease; lung or other visceral metastases are rare. Radiation therapy is the treatment of choice in patients with stage I disease and stage II disease where the nodes are<5 cm in maximum diameter. When a tumor contains both seminoma and nonseminoma components, patient management is directed by the more aggressive nonseminoma component.

TUMOR MARKERS

Careful monitoring of the serum tumor markersAFP andhCG is essential in the management of patients withGCT, as these markers are important for diagnosis, as prognostic indicators, in monitoring treatment response, and in the detection of early relapse. Approximately 70% of patients presenting with disseminated nonseminomatous GCT have increased serum concentrations of AFP and/or hCG. While hCG concentrations may be increased in patients with either nonseminoma or seminoma histology, the AFP concentration is increased only in patients with nonseminoma. The presence of an increased AFP level in a patient whose tumor showed only seminoma indicates that an occult nonseminomatous component exists and that the patient should be treated accordingly for nonseminomatous GCT. The serum lactate dehydrogenase (LDH) level serves as an additional marker of all GCTs, but it is not as specific as either AFP or hCG. LDH levels are increased in 50 to 60% patients with metastatic nonseminoma and in up to 80% of patients with advanced seminoma.

AFP,hCG, andLDHlevels should be determined before and after orchiectomy. Increased serum AFP and hCG concentrations decay according to first-order kinetics; the half-life is 24 to 36 h for hCG and 5 to 7 days for AFP. AFP and hCG should be assayed serially during and after treatment. The reappearance of hCG and/or AFP or the failure of these markers to decline according to the predicted half-life is an indicator of persistent or recurrent tumor.

TREATMENT

Stage I Nonseminoma If, after an orchiectomy (for clinical stage I disease), radiographs and physical examination show no evidence of disease, and serumAFP andhCGconcentrations either are normal or are declining to normal according to the known half-life, patients may be managed by either a nerve-sparing retroperitoneal lymph node dissection (RPLND) or surveillance. The retroperitoneal lymph nodes are pathologically involved byGCT(pathologic stage II) in 20 to 50% of these patients. The choice of surveillance or RPLND is based on the pathology of the primary tumor. If the primary tumor shows no pathologic evidence for lymphatic or vascular invasion *and* is limited to the testis (T1), then either option is reasonable. If lymphatic or vascular invasion is present *or* the tumor extends into the tunica, spermatic cord, or scrotum (T2 through T4), then surveillance should not be offered. Either approach should cure>95% of patients.

ARPLND is the standard operation for removal of the regional lymph nodes of the testis (retroperitoneal nodes). The operation removes the lymph nodes ipsilateral to the primary site and the nodal groups adjacent to the primary landing zone. The standard

(modified bilateral) RPLND removes all node-bearing tissue down to the bifurcation of the great vessels, including the ipsilateral iliac nodes. The major long-term effect of this operation is retrograde ejaculation and infertility. A nerve-sparing RPLND, usually accomplished by identification and dissection of individual nerve fibers, may avoid injury to the sympathetic nerves responsible for ejaculation. Normal ejaculation is preserved in approximately 90% of patients. Patients with pathologic stage I disease are observed, and only the 10% who relapse require additional therapy. If retroperitoneal nodes are found to be involved at RPLND, then a decision regarding adjuvant chemotherapy is made on the basis of the extent of retroperitoneal disease (see below).

Surveillance is an option in the management of clinical stage I disease when no vascular/lymphatic invasion is found and the primary tumor is classified as T1. Only 20 to 30% of patients have pathologic stage II disease, implying that mostRPLNDsin this situation are not therapeutic. Although surveillance has not been compared to RPLND in a randomized trial, all large studies show that surveillance and RPLND lead to equivalent long-term survival rates. Patient compliance is essential if surveillance is to be successful. Patients must be carefully followed with periodic chest radiography, physical examination,CT scan of the abdomen, and serum tumor marker determinations. The median time to relapse is about 7 months, and late relapses (later than 2 years) are rare. The 70 to 80% of patients who do not relapse require no intervention after orchiectomy; treatment is reserved for those who do relapse. When the primary tumor is classified as T2 through T4 *or* lymphatic/vascular invasion is identified, nerve-sparing RPLND is preferred. About 50% of these patients have pathologic stage II disease and are destined to relapse.

Stage II Nonseminoma Patients with limited, ipsilateral retroperitoneal adenopathy (nodes usually £3 cm in largest diameter) generally undergo a modified bilateralRPLND as primary management. Nearly all patients with pathologic stage II disease whose disease is completely resected by RPLND are cured. The local recurrence rate after a properly performed RPLND is very low. Depending on the extent of disease, the postoperative management options include either surveillance or two cycles of adjuvant chemotherapy. Surveillance is the preferred approach for patients with resected "low-volume" metastases (tumor nodes£2 cm in diameter, *and*<6 nodes are involved) because the probability of relapse is one-third or less. Because relapse occurs in³50% of patients with "high-volume" metastasis (>6 nodes involved, *or* any involved node >2 cm in largest diameter, *or* extranodal tumor extension), two cycles of adjuvant chemotherapy should be considered, as it results in cure in³98% of patients. Regimens consisting of etoposide (100 mg/m$_2$daily on days 1 through 5) plus cisplatin (20 mg/m$_2$daily on days 1 through 5) with or without bleomycin (30 units per day on days 2, 9, and 16) given at 3-week intervals are effective and well tolerated.

Stages I and II Seminoma Inguinal orchiectomy followed by retroperitoneal radiation therapy cures about 98% of patients with stage I seminoma. The dose of radiation (2500 to 3000 cGy) is low and well tolerated, and the in-field recurrence rate is negligible. About 2% of patients relapse with supradiaphragmatic or systemic disease. Surveillance has been proposed as an option, and studies have shown that about 15% of patients relapse. The median time to relapse is 12 to 15 months, and late relapses during surveillance (>5 years) may be more frequent than with nonseminoma. The relapse is usually treated with chemotherapy. Surveillance for clinical stage I seminoma

is generally not recommended.

Nonbulky retroperitoneal disease (stage IIA and IIB) is also treated with radiation therapy. Prophylactic supradiaphragmatic fields are not used. Relapses in the anterior mediastinum are unusual. Approximately 90% of patients achieve relapse-free survival with retroperitoneal masses<5 cm in diameter. Because at least one-third of patients with bulkier disease relapse, initial chemotherapy is preferred for stage IIC disease.

Chemotherapy for Advanced GCT Regardless of histology, patients with stage IIC and stage IIIGCT are treated with chemotherapy. Combination chemotherapy programs based on cisplatin at doses of 100 to 120 mg/m^2per cycle plus etoposide cure 70 to 80% of such patients, with or without bleomycin, depending on risk stratification (see below). A complete response (the complete disappearance of all clinical evidence of tumor on physical examination and radiography plus normal serum levels ofAFP andhCG for 1 month or more) occurs after chemotherapy alone in about 60% of patients, and another 10 to 20% become disease-free with surgical resection of all sites of residual disease. Lower doses of cisplatin result in inferior survival rates.

The toxicity of the cisplatin/bleomycin/etoposide (BEP) regimen may be substantial. Nausea, vomiting, and hair loss occur in most patients, although nausea and vomiting have been markedly ameliorated by modern antiemetic regimens. Myelosuppression is frequent, and symptomatic bleomycin pulmonary toxicity occurs in about 5% of patients. Treatment-induced mortality due to neutropenia with septicemia or bleomycin-induced pulmonary failure occurs in 1 to 3% of patients. Dose reductions for myelosuppression are rarely indicated. Long-term permanent toxicities include nephrotoxicity (reduced glomerular filtration and persistent magnesium wasting), ototoxicity, and peripheral neuropathy. When bleomycin is administered by weekly bolus injection, Raynaud's phenomenon appears in 5 to 10% of patients. Less often, other evidence of small blood vessel damage has been reported, including transient ischemic attacks and myocardial infarction.

Risk-Directed Chemotherapy Because not all patients are cured and treatment may cause significant toxicities, patients are stratified into "good-risk" and "poor-risk" groups according to pretreatment clinical features. For good-risk patients, the goal is to achieve maximum efficacy with minimal toxicity. For poor-risk patients, the goal is to identify more effective therapy with tolerable toxicity.

The International Germ Cell Cancer Consensus Group (IGCCCG) developed criteria to assign patients to three risk groups (good, intermediate, poor) (Table 96-2). Seminoma is either good or intermediate risk based on the absence or presence of nonpulmonary visceral metastases. Marker levels play no role in defining risk. No poor-risk category exists for seminoma. Nonseminomas have good-, intermediate-, and poor-risk categories based on the site of the primary tumor, the presence or absence of nonpulmonary visceral metastases, and marker levels.

For ~90% of patients with good-riskGCTs, four cycles of etoposide plus cisplatin (EP) or three cycles ofBEPproduce durable, complete responses, with minimal acute and chronic toxicity. Pulmonary toxicity is absent when bleomycin is not used and is rare when therapy is limited to 9 weeks; myelosuppression with neutropenic fever is less

frequent; and the treatment mortality rate is negligible. About 75% of intermediate-risk patients and 45% of poor-risk patients achieve durable complete remission with four cycles of BEP, and no regimen has proved superior. More effective therapy is needed.

Postchemotherapy Surgery Resection of residual metastases after the completion of chemotherapy is an integral part of therapy. If the initial histology is nonseminoma and the marker values have normalized, all sites of residual disease should be resected. In general, residual retroperitoneal disease requires a modified bilateralRPLND, which is associated with retrograde ejaculation. Thoracotomy (unilateral or bilateral) and neck dissection are less frequently required to remove residual mediastinal, pulmonary parenchymal, or cervical nodal disease. Viable tumor (seminoma, embryonal carcinoma, yolk sac tumor, or choriocarcinoma) will be present in 15%, mature teratoma in 40%, and necrotic debris and fibrosis in 45% of resected specimens. The frequency of teratoma or viable disease is highest in residual mediastinal tumors. If necrotic debris or mature teratoma is present, no further chemotherapy is necessary. If viable tumor is present but is completely excised, two additional cycles of chemotherapy are given.

If the initial histology is seminoma, mature teratoma is rarely present, and the most frequent finding is necrotic debris. For residual retroperitoneal disease, a completeRPLND is technically difficult owing to extensive postchemotherapy fibrosis. Observation is recommended when no radiographic abnormality exists or a residual mass<3 cm is present. Controversy exists over what to do when the residual mass exceeds 3 cm in diameter. About 25% of such masses contain viableGCT. Some investigators prefer excision or biopsy, but radiation therapy and surveillance are alternatives.

Salvage Chemotherapy Of patients with advancedGCT, 20 to 30% fail to achieve a durable complete response to first-line chemotherapy. A combination of cisplatin, ifosfamide and vinblastine (VeIP) will cure about 25% of patients as a second-line therapy. Patients are more likely to achieve a durable complete response toVeIP if they had a testicular primary tumor and relapsed from a prior complete remission to first-line cisplatin-containing chemotherapy. In contrast, if the patient failed to achieve a complete response or has a primary mediastinal nonseminoma, then VeIP is rarely beneficial. Those patients are candidates for dose-intensive treatment.

Chemotherapy consisting of dose-intensive, high-dose carboplatin (³1500 mg/m$_2$) plus etoposide (³1200 mg/m$_2$), with or without cyclophosphamide or ifosfamide, with peripheral blood stem cell support induces a complete response in 25 to 40% of patients who have progressed after ifosfamide-containing salvage chemotherapy. About one-half of the complete responses will be durable. High-dose therapy is the treatment of choice and standard of care for this patient population. Paclitaxel is active in previously treated patients and is being studied as a new component in conventional-dose and dose-intensive salvage therapy. Cure is still possible in some relapsed patients.

EXTRAGONADAL GCT AND MIDLINE CARCINOMA OF UNCERTAIN HISTOGENESIS

The prognosis and management of patients with extragonadalGCTsdepends on the

tumor histology and site of origin. All patients with a diagnosis of extragonadal GCT should have a testicular ultrasound examination. Nearly all patients with retroperitoneal or mediastinal seminoma achieve a durable complete response toBEP orEP. The clinical features of patients with primary retroperitoneal nonseminoma GCT are similar to those of patients with a primary of testis origin, and careful evaluation will find evidence of a primary testicular GCT in about two-thirds of cases. In contrast, a primary mediastinal nonseminomatous GCT is associated with a poor prognosis; one-third of patients are cured with standard therapy (four cycles of BEP). Patients with newly diagnosed mediastinal nonseminoma are considered to have poor-risk disease and should be considered for clinical trials testing regimens of possibly greater efficacy. In addition, mediastinal nonseminoma is associated with hematologic disorders, including acute myelogenous leukemia, myelodysplastic syndrome, and essential thrombocytosis unrelated to previous chemotherapy. These hematologic disorders are very refractory to treatment. Nonseminoma of any primary site may change into other malignant histologies such as embryonal rhabdomyosarcoma or adenocarcinoma. This is called malignant transformation. i(12p) has been identified in the transformed cell type, indicating GCT clonal origin.

A group of patients (most commonly men) with poorly differentiated tumors of unknown histogenesis, midline in distribution, and not associated with secretion ofAFP orhCG has been described; a few (10 to 20%) are cured by standard cisplatin-containing chemotherapy. i(12p) is present in about 25% of such tumors (the fraction that are cisplatin-responsive), confirming their origin from primitive germ cells. This finding is also predictive of the response to cisplatin-based chemotherapy and resulting long-term survival. These tumors are heterogeneous; neuroepithelial tumors and lymphoma may also present in this fashion.

FERTILITY

Infertility is an important consequence of the treatment ofGCTs. Preexisting infertility or impaired fertility is often present. Azoospermia and/or oligospermia are present at diagnosis in at least 50% of patients with testicular GCTs. Ejaculatory dysfunction is associated withRPLND, and germ cell damage may result from cisplatin-containing chemotherapy. Nerve-sparing techniques to preserve the retroperitoneal sympathetic nerves have made retrograde ejaculation less likely in the subgroups of patients who are candidates for this operation. Spermatogenesis does recur in some patients after chemotherapy. However, because of the significant risk of impaired reproductive capacity, semen analysis and cryopreservation of sperm in a sperm bank should be recommended to all patients before radiation therapy, chemotherapy, or RPLND.

(Bibliography omitted in Palm version)

Back to Table of Contents

97. GYNECOLOGIC MALIGNANCIES - *Robert C. Young*

OVARIAN CANCER

Incidence and Epidemiology Epithelial ovarian cancer is the leading cause of death from gynecologic cancer in the United States. In 2000, 23,100 new cases were diagnosed and 14,000 women died from ovarian cancer. The disease accounts for 5% of all cancer deaths in women in the United States; more women die of this disease than from cervical and endometrial cancer combined.

The age-specific incidence of the common epithelial type of ovarian cancer increases progressively and peaks in the eighth decade. Epithelial tumors, unlike germ cell and stromal tumors, are uncommon before the age of 40. Epidemiologic studies suggest higher incidences in industrialized nations and an association with disordered ovarian function, including infertility, nulliparity, frequent miscarriages, and use of ovulation-inducing drugs such as clomiphene. Each pregnancy reduces the ovarian cancer risk by about 10%, and breast feeding and tubal ligation also appear to reduce the risk. Oral contraceptives reduce the risk of ovarian cancer in patients with a familial history of cancer and in the general population. Many of these risk-reduction factors support the "incessant ovulation" hypothesis for ovarian cancer etiology, which implies that an aberrant repair process of the surface epithelium is central to ovarian cancer development. Estrogen replacement after menopause does not appear to increase the risk of ovarian cancer, although one study showed a modest increase in risk with >11 years of use.

Familial cases account for about 5% of all ovarian cancer, and a family history of ovarian cancer is a major risk factor. Compared to a lifetime risk of 1.6% in the general population, women with one affected first-degree relative have a 5% risk. In families with two or more affected first-degree relatives, the risk may exceed 50%. Three types of autosomal dominant familial cancer are recognized: (1) site-specific in which only ovarian cancer is seen, (2) families with cancer of the ovary and breast, and (3) the Lynch type II cancer family syndrome with nonpolyposis colorectal cancer, endometrial cancer, and ovarian cancer.

Etiology and Genetics In women with hereditary breast-ovarian cancer, two susceptibility loci have been identified: BRCA-1, located on chromosome 17q12-21, and BRCA-2, on 13q12-13. Both are tumor suppressor genes, and their protein products act as inhibitors of tumor growth. Both genes are large, and numerous mutations have been described; most are frameshift or nonsense mutations, and 86% produce truncated protein products. The implications of the many other mutations including many missense mutations are not known. The cumulative risk of ovarian cancer with critical mutations of BRCA-1 or -2 is 25%, compared to the lifetime risk of 50% for breast cancer for similar mutations. Men in such families have an increased risk of prostate cancer.

Cytogenetic analysis of sporadic epithelial ovarian cancers generally reveals complex karyotypic rearrangements. Structural abnormalities frequently appear on chromosomes 1 and 11, and loss of heterozygosity (LOH) is common on 3q, 6q, 11q, 13q, and 17. Abnormalities of oncogenes are frequently found in ovarian cancer and include c-*myc*,

H-*ras*, K-*ras*, and *neu.*

Ovarian tumors (usually not epithelial) are sometimes components of complex genetic syndromes. Peutz-Jeghers syndrome (mucocutaneous pigmentation and intestinal polyps) is associated with ovarian sex cord stromal tumors and Sertoli cell tumors in men. Patients with gonadal dysgenesis (46XY genotype or mosaic for Y-containing cell lines) develop gonadoblastomas, and women with nevoid basal cell carcinomas have an increased risk of ovarian fibromas.

Clinical Presentation and Differential Diagnosis Most patients with ovarian cancer are first diagnosed when the disease has already spread beyond the true pelvis. The occurrence of abdominal pain, bloating, and urinary symptoms usually indicates advanced disease. Localized ovarian cancer is generally asymptomatic. However, progressive enlargement of a localized ovarian tumor can produce urinary frequency or constipation, and rarely torsion of an ovarian mass causes acute abdominal pain or a surgical abdomen. In contrast to cervical or endometrial cancer, vaginal bleeding or discharge is rarely seen with early ovarian cancer. The diagnosis of early disease usually occurs with palpation of an asymptomatic adnexal mass during routine pelvic examination. However, most ovarian enlargements discovered this way, especially in premenopausal women, are benign functional cysts that characteristically resolve over one to three menstrual cycles. Adnexal masses in premenarchal or postmenopausal women are more likely to be pathologic. A solid, irregular, fixed pelvic mass is usually ovarian cancer. Other causes of adnexal masses include pedunculated uterine fibroids, endometriosis, benign ovarian neoplasms, and inflammatory lesions of the bowel.

Evaluation of patients with suspected ovarian cancer should include measurement of serum levels of the tumor marker CA-125. CA-125 determinants are glycoproteins with molecular masses from 220 to 1000 kDa, and a radioimmunoassay is used to determine circulating CA-125 antigen levels. Between 80 and 85% of patients with epithelial ovarian cancer have levels of CA-125 ³35 U/mL. Other malignant tumors can also elevate CA-125 levels, including cancers of the endometrium, cervix, fallopian tubes, pancreas, breast, lung, and colon. Certain nonmalignant conditions that can elevate CA-125 levels include pregnancy, endometriosis, pelvic inflammatory disease, and uterine fibroids. About 1% of normal females have serum CA-125 levels >35 U/mL. However, in postmenopausal women with an asymptomatic pelvic mass and CA-125 levels³65 U/mL, the test has a sensitivity of 97% and a specificity of 78%.

Screening In contrast to patients who present with advanced disease, patients with early ovarian cancers (stages I and II) are commonly curable with conventional therapy. Thus, effective screening procedures would improve the cure rate in this disease. Although pelvic examination can occasionally detect early disease, it is a relatively insensitive screening procedure. Transvaginal sonography has replaced the slower and less sensitive abdominal sonography, but significant false-positive results are noted, particularly in premenopausal women. In one study, 67 laparotomies were required to diagnose 1 primary ovarian cancer. Doppler flow imaging coupled with transvaginal ultrasound may improve accuracy and reduce the high rate of false positives.

CA-125 has been studied as a screening tool. Unfortunately, half of women with stages I and II ovarian cancer have CA-125 levels <65 U/mL. Other nonmalignant disorders

can elevate the CA-125 level, and both false-negative and -positive results have been high in most screening studies.

Attempts have been made to improve the sensitivity and specificity by combinations of procedures, commonly transvaginal ultrasound and CA-125 levels. In a screening study of 22,000 women, 42 had a positive screen and 11 had ovarian cancer (7 with advanced disease). In addition, eight women with a negative screen developed ovarian cancer. Thus, the false-positive rate would lead to a large number of unnecessary (i.e., negative) laparotomies if each positive screen resulted in a surgical exploration. The National Institutes of Health Consensus Conference recommended against screening for ovarian cancer among the general population without known risk factors for the disease. Although no evidence shows that screening saves lives, many physicians use annual pelvic examinations, transvaginal ultrasound, and CA-125 levels to screen women with a family history of ovarian cancer or breast/ovarian cancer syndromes.

Pathology Common epithelial tumors comprise most (85%) of the ovarian neoplasms. These may be benign (50%), frankly malignant (33%), or tumors of low malignant potential (16%) (tumors of borderline malignancy). Epithelial tumors of low malignant potential have the cytologic features of malignancy but do not invade the ovarian stroma. More than 75% of borderline malignancies present in early stage and generally occur in younger women. They have a much better natural history than their malignant counterpart.

There are five major subtypes of common epithelial tumors: serous (50%), mucinous (25%), endometroid (15%), clear cell (5%), and Brenner tumors (1%), the latter derived from the urothelium. Benign common epithelial tumors are almost always serous or mucinous and develop in women ages 20 to 60. They are frequently large (20 to 30 cm), bilateral, and cystic.

Malignant epithelial tumors are usually seen in women over 40. They present as solid masses, with areas of necrosis and hemorrhage. Masses>10 to 15 cm have usually already spread into the intraabdominal space. Spread eventually results in intraabdominal carcinomatosis, which leads to bowel and renal obstruction and cachexia.

Although most ovarian tumors are epithelial, two other important ovarian tumor types exist -- stromal and germ cell tumors. These tumors are distinct in their cell of origin but also have different clinical presentations and natural histories and are often managed differently (see below).

Metastasis to the ovary can occur from breast, colon, gastric, and pancreatic cancers, and the Krukenberg tumor was classically described as bilateral ovarian masses from metastatic mucin-secreting gastrointestinal cancers.

Staging and Prognostic Factors Laparotomy is often the primary procedure used to establish the diagnosis. Less invasive studies useful in defining the extent of spread include chest x-rays, abdominal computed tomography scans, and abdominal and pelvic sonography. If the woman has specific gastrointestinal symptoms, a barium enema or gastrointestinal series can be performed. Symptoms of bladder or renal dysfunction can

be evaluated by cystoscopy or intravenous pyelography.

A careful staging laparotomy will establish the stage and extent of disease and allow for the cytoreduction of tumor masses in patients with advanced disease. Proper laparotomy requires a vertical incision of sufficient length to ensure adequate examination of the abdominal contents. The presence, amount, and cytology of any ascites fluid should be noted. The primary tumor should be evaluated for rupture, excrescences, or dense adherence. Careful visual and manual inspection of the diaphragm and peritoneal surfaces is required. In addition to total abdominal hysterectomy and bilateral salpingo-oophorectomy, a partial omentectomy should be performed and the paracolic gutters inspected. Pelvic lymph nodes as well as para-aortic nodes in the region of the renal hilus should be biopsied. Since this surgical procedure defines stage, establishes prognosis, and determines the necessity for subsequent therapy, it should be performed by a surgeon with special expertise in ovarian cancer staging. Studies have shown that patients operated upon by gynecologic oncologists were properly staged 97% of the time, compared to 52 and 35% of cases staged by obstetricians/gynecologists and general surgeons, respectively. At the end of staging, 23% of women have stage I disease (cancer confined to the ovary or ovaries); 13% have stage II (disease confined to the true pelvis); 47% have stage III (disease spread into but confined to the abdomen); and 16% have stage IV disease (spread outside the pelvis and abdomen). The 5-year survival correlates with stage of disease: stage I -- 90%, stage II -- 70%, stage III -- 15 to 20%, and stage IV -- 1 to 5% (Table 97-1).

Prognosis in ovarian cancer is dependent not only upon stage but on the extent of residual disease and histologic grade. Patients presenting with advanced disease but left without significant residual disease after surgery have a median survival of 39 months, compared to 17 months for those with suboptimal tumor resection.

Prognosis of epithelial tumors is also highly influenced by histologic grade but less so by histologic type. In early-stage disease, survival is better in mucinous adenocarcinoma than endometrial and serous types, and clear cell carcinomas have the worst prognosis. Although grading systems differ among pathologists, all grading systems show a better prognosis for well- or moderately differentiated tumors and a poorer prognosis for poorly differentiated histologies. Typical 5-year survivals for patients with all stages of disease are: well differentiated -- 88%, moderately differentiated -- 58%, poorly differentiated -- 27%.

The prognostic significance of pre- and postoperative CA-125 levels is uncertain. Serum levels generally reflect volume of disease, and high levels usually indicate unresectability and a poorer survival. Postoperative levels, if elevated, usually indicate residual disease. Nevertheless, on multivariate analysis, CA-125 is not an independent prognostic factor because of the association with volume of disease. The rate of decline of CA-125 levels during initial therapy or the absolute level after one to three cycles of chemotherapy correlates with prognosis but is not sufficiently accurate to guide individual treatment decisions. Even when the CA-125 level falls to normal after surgery or chemotherapy, "second-look" laparotomy identifies residual disease in 60% of women. Other more quantitative approaches to define prognosis include ploidy analysis and image cytometry (automated analysis of cell morphology); they remain

investigational.

Genetic and biologic factors may influence prognosis. Increased tumor levels of p53 are associated with a worse prognosis in advanced disease. Epidermal growth factor receptors in ovarian cancer are associated with a high risk of progression, but the increased expression of HER-2/neu has given conflicting prognostic results, and expression of Mdr-1 has not been of prognostic value. HER-2/neu is being evaluated as a target for antibody therapy.

TREATMENT

The selection of therapy for patients with epithelial ovarian cancer depends upon the stage, extent of residual tumor, and histologic grade. In general, patients are considered in three separate treatment groups: (1) those with early (stages I and II) ovarian cancer and microscopic or no residual disease; (2) patients with advanced (stage III) disease but minimal residual tumor (<1 cm) after initial surgery; and (3) patients with bulky residual tumor and advanced (stage III or IV) disease.

Patients with stage I disease, no residual tumor, and well or moderately differentiated tumors need no adjuvant therapy after definitive surgery and 5-year survival exceeds 95%. For all other patients with early disease and those stage I patients with poor prognosis histologic grade, adjuvant therapy is probably warranted, and single-agent cisplatin or platinum-containing drug combinations used in advanced disease are appropriate. Five-year survival for this group exceeds 80%.

For the patients with advanced (stage III) disease but with limited or no residual disease after definitive cytoreductive surgery (about half of all stage III patients), the primary therapy is platinum-based combination chemotherapy. Approximately 70% of women respond to initial combination chemotherapy, and 40 to 50% have a complete regression of disease. Only about half of these patients are free of disease if surgically restaged. Although a variety of combinations are active, a randomized prospective trial of paclitaxel and cisplatin compared to cyclophosphamide and cisplatin in patients with more advanced disease demonstrated better results for the paclitaxel-cisplatin combination (response rate: 77 versus 64%; complete remission rate: 54 versus 33%, median survival: 37.5 versus 24.4 months). A subsequent trial of paclitaxel, 175 mg/m_2 by 3-h infusion, and carboplatin, dosed to an AUC (area under the curve) of 7.5, showed equal antitumor activity to paclitaxel plus cisplatin but substantially less toxicity.

Patients with advanced disease (stages III and IV) and bulky residual tumor are generally treated with a paclitaxel-platinum combination regimen as well and, while the overall prognosis is poorer, 5-year survival may reach 10 to 15%. In some instances, cytoreductive surgery can be performed after initial response to chemotherapy, and a multicenter European trial demonstrated that this strategy led to a significant improvement in progression-free interval and survival.

Historically, patients who had an excellent initial response to chemotherapy and have no clinical evidence of disease have had a second-look laparotomy. For patients with stage I ovarian cancer or for germ cell tumors, the operation rarely detects residual tumor and has been largely abandoned. Even for those with stages II and III epithelial tumors, the

second-look surgical procedure itself does not prolong overall survival. Its routine use cannot be recommended. Maintenance therapy does not prevent recurrences in patients in complete remission.

Patients with advanced disease whose disease recurs after initial treatment are usually not curable but may benefit significantly from limited surgery to relieve intestinal obstruction, localized radiation therapy to relieve pressure or pain from mass lesions or metastasis, or palliative chemotherapy. The selection of chemotherapy for palliation depends upon the initial regimen and evidence of drug resistance. Patients who have a complete regression of disease that lasts ³6 months respond to reinduction with the same agents. Patients relapsing within the first 6 months of initial therapy rarely do. Chemotherapeutic agents with>15% response rates in patients relapsing after initial combination chemotherapy include gemcitabine, topotecan, ifosfamide, etoposide, and hexamethylmelamine. Intraperitoneal chemotherapy (usually cisplatin) may be used if a small residual volume (<1 cm_3) of tumor exists. Progestational agents and antiestrogens produce responses in 5 to 15% of patients and have minimal side effects.

Borderline malignancy has a 95% 5-year survival even in stage III disease when managed with surgery. Radiation and chemotherapy are not useful.

OVARIAN GERM CELL TUMORS

Fewer than 5% of all ovarian tumors are germ cell in origin. They include teratoma, dysgerminoma, endodermal sinus tumor, and embryonal carcinoma. Germ cell tumors of the ovary generally occur in younger women (75% of ovarian malignancies in women<30), display an unusually aggressive natural history, and are commonly cured with less extensive nonsterilizing surgery and chemotherapy. Women cured of these malignancies are able to conceive and have normal children.

These neoplasms can be divided into three major groups: (1) benign tumors (usually dermoid cysts); (2) malignant tumors that arise from dermoid cysts; and (3) primitive malignant germ cell tumors including dysgerminoma, yolk sac tumors, immature teratomas, embryonal carcinomas, and choriocarcinoma.

Dermoid cysts are teratomatous cysts usually lined by epidermis and skin appendages. They often contain hair, and calcified bone or teeth can sometimes be seen on conventional pelvic x-ray. They are almost always curable by surgical resection. Approximately 1% of these tumors have malignant elements, usually squamous cell carcinoma.

Malignant germ cell tumors are usually large (median -- 16 cm). Bilateral disease is rare except in dysgerminoma (10 to 15% bilaterality). Abdominal or pelvic pain in young women is the usual presenting symptom. Serum human chorionic gonadoptropin (b-hCG) and afetoprotein levels are useful in the diagnosis and management of these patients. Before the advent of chemotherapy, extensive surgery was routine but has now been replaced by careful evaluation of extent of spread followed by resection of bulky disease and preservation of one ovary, uterus, and cervix, if feasible. This allows many affected women to preserve fertility. After surgical staging, 60 to 75% of women have stage I disease and 25 to 30% have stage III disease. Stages II and IV are

infrequent.

Most of the malignant germ cell tumors are managed with chemotherapy after surgery. Regimens used in testicular cancer such as PVB (cisplatin, vinblastine, bleomycin) and BEP (bleomycin 30 units IV weekly, etoposide 100 mg/m^2days 1 to 5, and cisplatin 20 mg/m^2days 1 to 5), with three or four courses given at 21-day intervals, have produced 95% long-term survival in patients with stages I to III disease. This regimen is the treatment of choice for all malignant germ cell tumors except grade I, stage I immature teratoma, where surgery alone is adequate, and perhaps early-stage dysgerminoma, where surgery and radiation therapy are used.

Dysgerminoma is the ovarian counterpart of testicular seminoma. The tumor is very sensitive to radiation therapy. The 5-year disease-free survival is 100% in early-stage patients and 61% in stage III disease. Unfortunately, the use of radiation therapy makes many patients infertile. BEP chemotherapy is equally or more effective and does not cause infertility. In incompletely resected patients with dysgerminoma, the 2-year disease-free survival was 95% and infertility was not observed. Combination chemotherapy (BEP) has replaced postoperative radiation therapy as the treatment of choice in women with ovarian dysgerminoma.

OVARIAN STROMAL TUMORS

Stromal tumors make up<10% of ovarian tumors. They are named for the stromal tissue involved: granulosa, theca, Sertoli, Leydig, and collagen-producing stromal cells. The granulosa and theca cell stromal cell tumors occur most frequently in the first three decades of life. Granulosa cell tumors frequently produce estrogen and cause menstrual abnormalities, bleeding, and precocious puberty. Endometrial carcinoma can be seen in 5% of these women, perhaps related to the persistent hyperestrogenism. Sertoli and Leydig cell tumors, when functional, produce androgens with resultant virilization or hirsutism. Some 75% of these stromal cell tumors present in stage I and can be cured with total abdominal hysterectomy and bilateral salpingo-oophorectomy. Stromal tumors generally grow slowly, and recurrences can occur 5 to 10 years after initial surgery. Neither radiation therapy nor chemotherapy have been documented to be consistently effective, and surgical management remains the primary treatment.

CARCINOMA OF THE FALLOPIAN TUBE

The fallopian tube is the least common site of cancer in the female genital tract although its epithelial surface far exceeds that of the ovary, where epithelial cancer is 20 times more common. Approximately 300 new cases occur yearly; 90% are papillary serous adenocarcinomas, with the remainder being mixed mesodermal, endometroid, and transitional cell tumors. BRCA-1 and -2 mutations are found in 7% of cases. The gross and microscopic characteristics and the spread of the tumor are similar to those of ovarian cancer but can be distinguished if the tumor arises from the endosalpinx, the tubal epithelium shows a transition between benign and malignant, and the ovaries and endometrium are normal or minimally involved. The differential diagnosis includes primary or metastatic ovarian cancer, chronic salpingitis, tuberculous salpingitis, salpingitis isthmica nodosa, and cautery artifact.

Unlike patients with ovarian cancer, patients frequently present with early symptoms, usually postmenopausal vaginal bleeding, pain, and leukorrhea. Surgical staging is similar to that used for ovarian cancer, and prognosis is related to stage and extent of residual disease. Patients with stages I and II disease are generally treated with surgery alone or with surgery and pelvic radiation therapy, although radiation therapy does not clearly improve 5-year survival (5-year survival stage I: 74 versus 75%, stage II: 43 versus 48%). Patients with stages III and IV disease are treated with the same chemotherapy regimens used in advanced ovarian carcinoma, and 5-year survival is similar (stage III -- 20%, stage IV -- 5%).

UTERINE CANCER

Carcinoma of the endometrium is the most common female pelvic malignancy. Approximately 36,100 new cases are diagnosed yearly, although in most (75%) tumor is confined to the uterine corpus at diagnosis and therefore most can be cured. The 6500 deaths yearly make uterine cancer only the seventh leading cause of cancer death in females. It is primarily a disease of postmenopausal women, although 25% of cases occur in women <age 50 and 5%<age 40. The disease is common in Eastern Europe and the United States and uncommon in Asia.

Phenotypic characteristics and risk factors common in patients with endometrial cancer include obesity, altered menstruation, low fertility index, late menopause, anovulation, and postmenopausal bleeding. Exposure to unopposed estrogen from either endogenous or exogenous sources may play a central etiologic role. Women taking tamoxifen for breast cancer treatment or prevention have a twofold increased risk.

Endometrial carcinoma occurs most often in the sixth and seventh decades of life. Symptoms often include abnormal vaginal discharge (90%); abnormal bleeding (80%), which is usually postmenopausal; and leukorrhea (10%). Evaluation of such patients should include a history and physical and pelvic examinations followed by an endometrial biopsy or a fractional dilation and curettage. Outpatient procedures such as endometrial biopsy or aspiration curettage can be used but are definitive only when positive.

Between 75 and 80% of all endometrial carcinomas are adenocarcinomas, and the prognosis depends upon stage, histologic grade, and extent of myometrial invasion. Grade I tumors are highly differentiated adenocarcinomas, grade II contain some solid areas, and grade III tumors are largely solid or undifferentiated. Adenocarcinoma with squamous differentiation is seen in 10% of patients; the most differentiated form is known as *adenoacanthoma*, and the poorly differentiated form is called *adenosquamous carcinoma*. Other less common pathologies include mucinous carcinoma (5%) and papillary serous carcinoma (<10%). This latter type has a natural history similar to ovarian carcinoma and should be managed as an ovarian cancer. Rarer histologies include secretory (2%), ciliated, clear cell, and undifferentiated carcinomas.

The staging of endometrial cancer requires surgery to establish the extent of disease and the depth of myometrial invasion. Peritoneal fluid should be sampled; the abdomen and pelvis explored; and pelvic and para-aortic lymphadenectomy performed depending

upon the histology, grade, and depth of invasion in the uterine specimen on frozen section. After evaluation and staging, 74% of patients are stage I, 13% are stage II, 9% are stage III, and 3% are stage IV. Five-year survival by stage is as follows: stage I -- 89%, stage II -- 80%, stage III -- 30%, and stage IV -- 9% (Table 97-1).

Patients with uncomplicated endometrial carcinoma are effectively managed with total abdominal hysterectomy and bilateral salpingo-oophorectomy. Pre- or postoperative irradiation has been used, and although vaginal cuff recurrence is reduced, survival is not altered. In women with poor histologic grade, deep myometrial invasion, or extensive involvement of the lower uterine segment or cervix, intracavitary or external beam irradiation is warranted.

About 15% of women have endometrial carcinoma with extension to the cervix only (stage II), and management depends upon the extent of cervical invasion. Superficial cervical invasion can be managed like stage I disease, but extensive cervical invasion requires radical hysterectomy or preoperative radiotherapy followed by extrafascial hysterectomy. Once disease is outside the uterus but still confined to the true pelvis (stage III), management generally includes surgery and irradiation. Patients who have involvement only of the ovary or fallopian tubes generally do well with such therapy (5-year survivals of 80%). Other stage III patients with disease extending beyond the adnexa or those with serous carcinomas of the endometrium have a significantly poorer prognosis (5-year survival of 15%).

Patients with stage IV disease (outside the abdomen or invading the bladder or rectum) are treated palliatively with irradiation, surgery, and/or progestational agents. Progestational agents produce responses in about 25% of patients. Well-differentiated tumors respond most frequently, and response can be correlated with the level of progesterone receptor expression in the tumor. The commonly used progestational agents hydroxyprogesterone (Dilalutin), megastrol (Megace), and deoxyprogesterone (Provera) all produce similar response rates, and the antiestrogen tamoxifen (Nolvadex) produces responses in 10 to 25% of patients in a salvage setting.

Chemotherapy is not very successful in advanced endometrial carcinoma. The most active single agents with consistent response rates of ^{3}20% include cisplatin, carboplatin, doxorubicin, epirubicin, and paclitaxel. Combinations of drugs with or without progestational agents have generally produced response rates similar to single agents.

CERVIX CANCER

Carcinoma of the cervix was once the most common cause of cancer death in women, but over the past 30 years, the mortality rate has decreased by 50% due to widespread screening with the Pap smear. Cervix cancer trails breast, lung, colorectal, endometrium, and ovarian cancers in incidence. In 2000, ~12,800 new cases of invasive cervix cancer occurred, and >50,000 cases of carcinoma in situ were detected. There were 4600 deaths from the disease, and of those patients, ~85% had never had a Pap smear. It remains the major gynecologic cancer in underdeveloped countries. It is more common in lower socioeconomic groups, in women with early initial sexual activity and/or multiple sexual partners, and in smokers. Venereal transmission of human

papilloma virus (HPV) has an important etiologic role. Over 66 types of HPVs have been isolated, and many are associated with genital warts. Those types associated with cervical carcinoma are 16, 18, 31, 45, and 51 to 53. These, along with many other types, are also associated with cervical intraepithelial neoplasia (CIN). The protein product of HPV-16, the E7 protein, binds and inactivates the tumor suppressor gene Rb, and the E6 protein of HPV-18 has sequence homology to the SV40 large T antigen and has the capacity to bind and inactivate the tumor suppressor gene p53. E6 and E7 are both necessary and sufficient to cause cell transformation in vitro. These binding and inactivation events may explain the carcinogenic effects of the viruses (Chap. 188).

UncomplicatedHPVlower genital tract infection and condylomatous atypia of the cervix can progress toCIN. This lesion precedes invasive cervical carcinoma and is classified as low-grade squamous intraepithelial lesion (SIL), high-grade SIL, and carcinoma in situ. Carcinoma in situ demonstrates cytologic evidence of neoplasia without invasion through the basement membrane, can persist unchanged for 10 to 20 years, but eventually progresses to invasive carcinoma.

The Pap smear is 90 to 95% accurate in detecting early lesions such asCIN but is less sensitive in detecting cancer when frankly invasive cancer or fungating masses are present. Inflammation, necrosis, and hemorrhage may produce false-positive smears, and colposcopic-directed biopsy is required when any lesion is visible on the cervix, regardless of Pap smear findings. The American Cancer Society recommends that women after onset of sexual activity, or>age 20, have two consecutive yearly smears. If negative, smears should be repeated every 3 years. The American College of Obstetrics and Gynecology recommends yearly Pap smears with routine annual pelvic and breast examinations. The Pap smear can be reported as normal (includes benign, reactive or reparative changes); atypical squamous cells of undetermined significance (ASCUS); low- or high-grade CIN; or frankly malignant. Women with ASCUS or low-grade CIN should have repeat smears in 3 to 6 months and be tested forHPV. Women with high-grade CIN or frankly malignant Pap smears should have colposcopic-directed cervical biopsy. Colposcopy is a technique using a binocular microscope and 3% acetic acid applied to the cervix in which abnormal areas appear white and can be biopsied directly. Cone biopsy is still required when endocervical tumor is suspected, colposcopy is inadequate, the biopsy shows microinvasive carcinoma, or when a discrepancy is noted between the Pap smear and the colposcopic findings. Cone biopsy alone is therapeutic for CIN in many patients, although a less radical electrocautery excision may be sufficient.

Approximately 80% of invasive cervix carcinomas are squamous cell tumors, 10 to 15% are adenocarcinomas, 2 to 5% are adenosquamous with epithelial and glandular structures, and 1 to 2% are clear cell mesonephric tumors.

Patients with cervix cancer generally present with abnormal bleeding or postcoital spotting that may increase to intermenstrual or prominent menstrual bleeding. Yellowish vaginal discharge, lumbosacral back pain, and urinary symptoms can also be seen.

The staging of cervical carcinoma is clinical and generally completed with a pelvic examination under anesthesia with cystoscopy and proctoscopy. Chest x-rays, intravenous pyelograms, and computed tomography are generally required, and

magnetic resonance imaging (MRI) may be used to assess extracervical extension. Stage 0 is carcinoma in situ, stage I is disease confined to the cervix, stage II disease invades beyond the cervix but not to the pelvic wall or lower third of the vagina, stage III disease extends to the pelvic wall or lower third of the vagina or causes hydronephrosis, stage IV is present when the tumor invades the mucosa of bladder or rectum or extends beyond the true pelvis. Five-year survivals are as follows: stage I -- 85%, stage II -- 60%, stage III -- 33%, and stage IV -- 7% (Table 97-1).

Carcinoma in situ (stage 0) can be managed successfully by cone biopsy or by abdominal hysterectomy. For stage I disease, results appear equivalent for either radical hysterectomy or radiation therapy. Patients with stages II to IV disease are primarily managed with radical radiation therapy or combined modality therapy. Retroperitoneal lymphadenectomy has no proven therapeutic role. Pelvic exenterations, although uncommon, are performed for centrally recurrent or persistent disease. Advances have been made in the reconstruction of the vagina, bladder, and rectum following this operation.

In women with locally advanced disease (stages IIB to IVA), platinum-based chemotherapy given concomitantly with radiation therapy improves survival compared to radiation therapy alone. Cisplatin, 75 mg/m$_2$over 4 h, followed by 5-fluorouracil (5-FU) 4 g given by 96-h infusion on days 1 to 5 of radiation therapy, is a common regimen. Two additional cycles of chemotherapy are given at 3-week intervals. Concurrent chemoradiotherapy reduced the risk of recurrence by 30 to 50% across wide spectrum of stages and presentations and is the treatment of choice in stages IIB to IV cervix cancer.

Chemotherapy has been used in patients with unresectable advanced disease or recurrent disease. Active agents with^{3}20% response rates include cisplatin,5-FU, ifosfamide, and irinotecan. No combination of agents has proved better than single agents. Intraarterial chemotherapy has been studied, either pre-or postoperatively, but is associated with substantial local toxicity and response rates of 20%.

GESTATIONAL TROPHOBLASTIC NEOPLASIA

Gestational choriocarcinoma accounts for <1% of female gynecologic malignancies and can be cured with appropriate chemotherapy. Deaths from this disease have become rare in the United States. The spectrum of disease ranges from benign hydatidiform mole to trophoblastic malignancy (placental-site trophoblastic tumor and choriocarcinoma).

Epidemiology In the United States, the incidence is about 1 per 1000 pregnancies; in Asia, 2 per 1000 pregnancies. Maternal age >45 years is a risk factor for hydatidiform mole. A prior history of molar pregnancy is also a risk factor. Choriocarcinoma occurs approximately once in 25,000 pregnancies or once in 20,000 live births. Prior history of hydatidiform mole is a risk factor for choriocarcinoma. A woman with a molar pregnancy is 1000 times more likely to develop choriocarcinoma than a woman with a prior normal-term pregnancy.

Pathology and Etiology The trophoblastic neoplasms have been divided by

morphology into complete or partial hydatidiform mole, invasive mole, placental-site trophoblastomas, and choriocarcinomas. Hydatidiform moles contain clusters of villi with hydropic changes, hyperplasia of the trophoblast, and the absence of fetal vessels. Invasive moles differ only by invasion into the uterine myometrium. Placental-site trophoblastic tumors are predominately made up of cytotrophoblast cells arising from the placental implantation site. Choriocarcinomas consist of anaplastic trophoblastic tissue with both cytotrophoblastic and syncytiotrophoblastic elements and no identifiable villi.

Complete moles result from uniparental disomy in which loss of the maternal genes (23 autosomes plus X) occurs by unknown mechanisms and is followed by duplication of the paternal haploid genome (23 autosomes plus X). Uncommonly (5%), moles result from dispermic fertilization of an empty egg, resulting in either 46XY or 46XX genotype. Partial moles result from dispermic fertilization of an egg with retention of the maternal haploid set of chromosomes, resulting in diandric triploidy (Chap. 65).

Clinical Presentation Molar pregnancies are generally associated with first-trimester bleeding, ectopic pregnancies, or threatened abortions. The uterus is inappropriately large for the length of gestation, andb-hCGlevels are higher than expected. Fetal parts and heart sounds are not present. The diagnosis is generally made by the passage of grapelike clusters from the uterus, but ultrasound demonstration of the hydropic mole can be diagnostic. Patients suspected of a molar pregnancy require a chest film, careful pelvic examinations, and weekly serial monitoring of b-hCG levels.

TREATMENT

Patients with hydatidiform moles require surgical evacuation coupled with postevacuation monitoring ofb-hCGlevels. In most women (80%), the b-hCG titer progressively declines within 8 to 10 days of evacuation (serum half-life is 24 to 36 h). Patients should be monitored on a monthly basis and should not become pregnant for at least a year. Patients found to have invasive mole at curettage are generally treated with hysterectomy and chemotherapy. Approximately half of patients with choriocarcinoma develop the malignancy after a molar pregnancy, and the other half develop the malignancy after abortion, ectopic pregnancy, or occasionally after a normal full-term pregnancy.

Chemotherapy is generally used for gestational trophoblastic neoplasia and is often used in hydatidiform mole ifb-hCGlevels rise or plateau or if metastases develop. Patients with invasive mole or choriocarcinoma require chemotherapy. Several regimens are effective, including methotrexate at 30 mg/m^{2}intramuscularly on a weekly basis until b-hCG titers are normal. However, methotrexate (1 mg/kg) every other day for 4 days followed by leukovorin (0.1 mg/kg) intravenously 24 h after methotrexate is associated with a cure rate of ^{3}90% and low toxicity. Intermittent courses are continued until the b-hCG titer becomes undetectable for 3 consecutive weeks, and then patients are monitored monthly for a year.

Patients with high-risk tumors (highb-hCGlevels, disease presenting ^{3}4 months after antecedent pregnancy, brain or liver metastasis, or failure of single-agent methotrexate) are initially treated with combination chemotherapy. MAC chemotherapy with

methotrexate, actinomycin-D, and cyclophosphamide has been the most commonly used regimen, with cycles of therapy given every 3 weeks until complete remission. Other effective regimens include EMA-CO (a cyclic non-cross-resistant combination of etoposide, methotrexate, and dactinomycin alternating with cyclophosphamide and vincristine); cisplatin, bleomycin, and vinblastine; and cisplatin, etoposide, and bleomycin. EMA-CO is now the regimen of choice for patients with high-risk disease because of excellent survival rates (>80%) and less toxicity than MAC. The use of etoposide carries a 1.5% lifetime risk of acute myeloid leukemia (16-fold relative risk). Because of this problem, etoposide-containing regimens should be reserved for patients with high risk features. Patients with brain or liver metastasis are usually treated with local irradiation to metastatic sites in conjunction with chemotherapy. Long-term studies of patients cured of trophoblastic disease have not demonstrated an increased risk of maternal complications or fetal abnormalities with subsequent pregnancies.

(Bibliography omitted in Palm version)

Back to Table of Contents

98. SOFT TISSUE AND BONE SARCOMAS AND BONE METASTASES -

Shreyaskumar R. Patel*, *Robert S. Benjamin

Sarcomas are rare (less than 1% of all malignancies) mesenchymal neoplasms that arise in bone and soft tissues. These tumors are usually of mesodermal origin, although a few are derived from neuroectoderm, and they are biologically distinct from the more common epithelial malignancies. Sarcomas affect all age groups; 15% are found in children younger than age 15, and 40% occur after age 55. Sarcomas are one of the most common solid tumors of childhood and are the fifth most common cause of cancer deaths in children. Sarcomas may be divided into two groups, those derived from bone and those derived from soft tissues.

SOFT TISSUE SARCOMAS

Soft tissues include muscles, tendons, fat, fibrous tissue, synovial tissue, vessels, and nerves. Approximately 60% of soft tissue sarcomas arise in the extremities, with the lower extremities involved three times as often as the upper extremities. Thirty percent arise in the trunk, the retroperitoneum accounting for 40% of all trunk lesions. The remaining 10% arise in the head and neck.

INCIDENCE

Approximately 7800 new cases of soft tissue sarcomas occurred in the United States in 1999. The annual age-adjusted incidence is approximately 2 per 100,000 population, but the incidence varies with age. Soft tissue sarcomas constitute 0.7% of all cancers in the general population and 6.5% of all cancers in children.

EPIDEMIOLOGY

Malignant transformation of a benign soft tissue tumor is extremely rare, with the exception that malignant peripheral nerve sheath tumors (neurofibrosarcoma, malignant schwannoma) can arise from neurofibromas in patients with neurofibromatosis. Several etiologic factors have been implicated in soft tissue sarcomas.

Environmental Factors Trauma or previous injury is rarely involved, but sarcomas can arise in scar tissue resulting from a prior operation, burn, fracture, or foreign body implantation. Chemical carcinogens such as polycyclic hydrocarbons, asbestos, and dioxin may be involved in the pathogenesis.

Iatrogenic Factors Sarcomas in bone or soft tissues occur in patients who are treated with radiation therapy. The tumor nearly always arises in the irradiated field. The risk increases with time.

Viruses Kaposi's sarcoma (KS) in patients with HIV type 1, classic KS, and KS in HIV-negative homosexual men is caused by human herpes virus (HHV8) (Chap. 185). No other sarcomas are associated with viruses.

Immunologic Factors Congenital or acquired immunodeficiency, including therapeutic immunosuppression, increases risk of sarcoma.

Genetic Factors Li-Fraumeni syndrome is a familial cancer syndrome in which affected individuals have germ-line abnormalities of the tumor suppressor gene *p53* and an increased incidence of soft tissue sarcomas and other malignancies, including breast cancer, osteosarcoma, brain tumor, leukemia, and adrenal carcinoma (Chap. 81). Neurofibromatosis 1 (NF-1, peripheral form, von Recklinghausen's disease) is characterized by multiple neurofibromas and cafe au lait spots. Neurofibromas occasionally undergo malignant degeneration to become malignant peripheral nerve sheath tumors. The gene for NF-1 is located in the pericentromeric region of chromosome 17 and encodes neurofibromin, a tumor suppressor protein with GTPase-activating activity that inhibits Ras function (Chap. 370). Germ-line mutation of the *Rb-1* locus (chromosome 13q14) in patients with inherited retinoblastoma is associated with the development of osteosarcoma in those who survive the retinoblastoma and of soft tissue sarcomas unrelated to radiation therapy. Other soft tissue tumors, including desmoid tumors, lipomas, leiomyomas, neuroblastomas, and paragangliomas, occasionally show a familial predisposition.

Ninety percent of synovial sarcomas contain a characteristic chromosomal translocation t(X;18) (p11;q11) involving a nuclear transcription factor on chromosome 18 called *SYT* and two breakpoints on X. Patients with translocations to the second X breakpoint (*SSX2*) may have longer survival than those with translocations involving *SSX1*.

Insulin-like growth factor (IGF) type 2 is produced by some sarcomas and may act as an autocrine growth factor and as a motility factor that promotes metastatic spread. IGF-2 stimulates growth through IGF-1 receptors but its effects on motility are through different receptors. If secreted in large amounts, IGF-2 may produce hypoglycemia (Chaps. 100 and334).

CLASSIFICATION

Approximately 20 different groups of sarcomas are recognized on the basis of the pattern of differentiation toward normal tissue. For example, rhabdomyosarcoma shows evidence of skeletal muscle fibers with cross-striations; leiomyosarcomas contain interlacing fascicles of spindle cells resembling smooth muscle; and liposarcomas contain adipocytes. When precise characterization of the group is not possible, the tumors are called *unclassified sarcomas*. All of the primary bone sarcomas also can arise from soft tissues (e.g., extraskeletal osteosarcoma). The entity *malignant fibrous histiocytoma* includes many tumors previously classified as fibrosarcomas or as pleomorphic variants of other sarcomas and is characterized by a mixture of spindle (fibrous) cells and round (histiocytic) cells arranged in a storiform pattern with frequent giant cells and areas of pleomorphism.

For purposes of treatment, most soft tissue sarcomas can be considered together. However, some specific tumors have distinct features. For example, *liposarcoma* can have a spectrum of behaviors. Pleomorphic liposarcomas and dedifferentiated liposarcomas behave like other high-grade sarcomas; in contrast, well-differentiated liposarcomas (better termed *atypical lipomatous tumors*) lack metastatic potential, and myxoid liposarcomas metastasize infrequently but, when they do, have a predilection for unusual metastatic sites containing fat, such as the retroperitoneum, mediastinum, and

subcutaneous tissue. Rhabdomyosarcomas, Ewing's sarcoma, and other small cell sarcomas tend to be more aggressive, and are more responsive to chemotherapy than other soft tissue sarcomas.

DIAGNOSIS

The most common presentation is an asymptomatic mass. Mechanical symptoms referable to pressure, traction, or entrapment of nerves or muscles may be present. All new and persistent or growing masses should be biopsied, either by a cutting needle (core-needle biopsy) or by a small incision, placed so that it can be encompassed in the subsequent excision without compromising a definitive resection. Sarcomas tend to metastasize through the blood rather than the lymphatic system; lymph node metastases occur in 5%, except in synovial and epithelioid sarcomas, clear-cell sarcoma (melanoma of the soft parts), angiosarcoma, and rhabdomyosarcoma where nodal spread may be seen in 17%. The pulmonary parenchyma is the most common site of metastases. Exceptions are leiomyosarcomas arising in the gastrointestinal tract, which metastasize to the liver; myxoid liposarcomas, which seek fatty tissue: and clear-cell sarcomas, which may metastasize to bones. Central nervous system metastases are rare, except in alveolar soft part sarcoma.

Radiographic Evaluation Imaging of the primary tumor is best with plain radiographs and magnetic resonance imaging (MRI) for tumors of the extremities or head and neck and by computed tomography (CT) for tumors of the chest, abdomen, or retroperitoneal cavity. A radiograph and CT scan of the chest are important for the detection of lung metastases. Other imaging studies may be indicated, depending on the symptoms, signs, or histology.

STAGING AND PROGNOSIS

The histologic grade, relationship to fascial planes, and size of the primary tumor are the most important prognostic factors. The newly revised American Joint Commission on Cancer (AJCC) staging system is shown in Table 98-1. Prognosis is related to the stage. Cure is common in the absence of metastatic disease, but a small number of patients with metastases can also be cured. Most patients with stage IV disease die within 6 to 12 months, but some patients may live with slowly progressive disease for many years.

TREATMENT

AJCC stage I patients are adequately treated with surgery alone. Stage II patients require adjuvant radiation therapy. Stage III patients require adjuvant chemotherapy. Stage IV patients are managed primarily with chemotherapy with or without other modalities.

Surgery Soft tissue sarcomas tend to grow along fascial planes, with the surrounding soft tissues compressed to form a pseudocapsule that gives the sarcoma the appearance of a well-encapsulated lesion. This is invariably deceptive, because "shelling out" or marginal excision of such lesions results in a 50 to 90% probability of local recurrence. Wide excision with a negative margin, incorporating the biopsy site, is the standard surgical procedure for local disease. The adjuvant use of radiation therapy

and/or chemotherapy improves the local control rate and permits the use of limb-sparing surgery with a local control rate (85 to 90%) comparable to that achieved by radical excisions and amputations. Limb-sparing approaches are indicated except when negative margins are not obtainable, when the risks of radiation are prohibitive, or when neurovascular structures are involved so that resection will result in serious functional consequences to the limb.

Radiation Therapy External beam radiation therapy is an adjuvant to limb-sparing surgery for improved local control. Preoperative radiation therapy allows the use of smaller fields and smaller doses but results in a higher rate of wound complications. Postoperative radiation therapy must be given to larger fields, as the entire surgical bed must be encompassed, and in higher doses to compensate for hypoxia in the operated field. Brachytherapy or interstitial therapy, in which the radiation source is inserted into the tumor bed, is comparable in efficacy (except in low grade lesions), less time consuming, and less expensive.

Adjuvant Chemotherapy Chemotherapy is the mainstay of treatment for Ewing's/peripheral neuroepithelial tumors (PNET) and rhabdomyosarcomas. Meta-analysis of 14 randomized trials revealed a highly significant improvement in local control and disease-free survival in favor of doxorubicin-based chemotherapy. Overall survival is improved only for extremity sarcomas, however. An alternative approach is to treat such patients preoperatively with chemotherapy (neoadjuvant therapy); the subset of patients who respond continue adjuvant therapy postoperatively, and the nonresponders can be spared the toxicity of systemic therapy to which they are unlikely to respond. Neither strategy has been proved superior.

Advanced Disease Metastatic soft tissue sarcomas are largely incurable, but up to 20% of patients who achieve a complete response become long-term survivors. The therapeutic intent, therefore, is to produce a complete remission with chemotherapy and/or surgery. Surgical resection of metastases, whenever possible, is an integral part of the management. Some patients benefit from repeated surgical excision of metastases. Despite their histologic heterogeneity, the sensitivity to chemotherapy of most soft tissue sarcomas is poor. The two most active chemotherapeutic agents are doxorubicin and ifosfamide. There is a steep dose-response relationship for these drugs in sarcomas. Dacarbazine (DTIC) in combination with doxorubicin may be more active than the single agents. Vincristine, etoposide, and dactinomycin are effective in Ewing's sarcoma and rhabdomyosarcoma, especially in children. Chondrosarcomas and leiomyosarcomas arising from the gastrointestinal tract are unresponsive to standard chemotherapeutic drugs.

BONE SARCOMAS

INCIDENCE AND EPIDEMIOLOGY

Bone sarcomas are rarer than soft tissue sarcomas; they accounted for only 0.2% of all new malignancies and approximately 2500 new cases in the United States in 1999. Several benign bone lesions have the potential for malignant transformation. Enchondromas and osteochondromas can transform into chondrosarcoma; fibrous dysplasia, bone infarcts, and Paget's disease of bone can transform into either

malignant fibrous histiocytoma or osteosarcoma.

CLASSIFICATION

Benign Tumors The common benign bone tumors include enchondroma, osteochondroma, chondroblastoma, and chondromyxoid fibroma, of cartilage origin; osteoid osteoma and osteoblastoma, of bone origin; fibroma and desmoplastic fibroma, of fibrous tissue origin; hemangioma, of vascular origin; and giant cell tumor, of unknown origin.

Malignant Tumors The most common malignant tumors of bone are plasma cell tumors (Chap. 113). The four most common malignant nonhematopoietic bone tumors are osteosarcoma, chondrosarcoma, Ewing's sarcoma, and malignant fibrous histiocytoma. Rare malignant tumors include chordoma (of notochordal origin), malignant giant cell tumor and adamantinoma (of unknown origin), and hemangioendothelioma (of vascular origin).

Musculoskeletal Tumor Society Staging System Sarcomas of bone are staged according to the Musculoskeletal Tumor Society staging system based on grade and compartmental localization. A Roman numeral reflects the tumor grade: stage I is low-grade, stage II is high-grade, and stage III includes tumors of any grade that have lymph node or distant metastases. In addition, the tumor is given a letter reflecting its compartmental localization. Tumors designated A are intracompartmental (i.e., confined to the same soft tissue compartment as the initial tumor), and tumors designated B are extracompartmental (i.e., extending into the adjacent soft tissue compartment or into bone).

OSTEOSARCOMA

Osteosarcoma, accounting for almost 45% of all bone sarcomas, is a spindle cell neoplasm that produces osteoid (unmineralized bone) or bone. About 60% of all osteosarcomas occur in children and adolescents in the second decade of life, and about 10% occur in the third decade of life. Osteosarcomas in the fifth and sixth decades of life are frequently secondary to either radiation therapy or transformation in a preexisting benign condition, such as Paget's disease. Males are affected 1.5 to 2 times as often as females. Osteosarcoma has a predilection for metaphyses of long bones; the most common sites of involvement are the distal femur, proximal tibia, and proximal humerus. The classification of osteosarcoma is complex, but 75% of osteosarcomas fall in the "classic" category, which include osteoblastic, chondroblastic, and fibroblastic osteosarcomas. The remaining 25% are classified as "variants" on the basis of (1) clinical characteristics, as in the case of osteosarcoma of the jaw, postradiation osteosarcoma, or Paget's osteosarcoma; (2) morphologic characteristics, as in the case of telangiectatic osteosarcoma, small cell osteosarcoma, or epithelioid osteosarcoma; or (3) location, as in parosteal or periosteal osteosarcoma. Diagnosis usually requires a synthesis of clinical, radiologic, and pathologic features. Patients typically present with pain and swelling of the affected area. A plain radiograph reveals a destructive lesion with a moth-eaten appearance, a spiculated periosteal reaction (sunburst appearance), and a cuff of periosteal new bone formation at the margin of the soft tissue mass (Codman's triangle). ACT scan of the primary tumor is best for defining

bone destruction and the pattern of calcification, whereasMRI is better for defining intramedullary and soft tissue extension. A chest radiograph and CT scan are used to detect lung metastases. Metastases to the bony skeleton should be imaged by a bone scan. Almost all osteosarcomas are hypervascular. Angiography is not helpful for diagnosis, but it is the most sensitive test for assessing the response to preoperative chemotherapy. Pathologic diagnosis is established either with a core-needle biopsy, where feasible, or with an open biopsy with an appropriately placed incision that does not compromise future limb-sparing resection. Most osteosarcomas are high-grade. The most important prognostic factor for long-term survival is response to chemotherapy. Preoperative chemotherapy followed by limb-sparing surgery (which can be accomplished in >80% of patients) followed by postoperative chemotherapy is standard management. The effective drugs are doxorubicin, ifosfamide, cisplatin, and high-dose methotrexate with leucovorin rescue. The various combinations of these agents that have been used have all been about equally successful. Long-term survival rates in extremity osteosarcoma range from 60 to 80%. Osteosarcoma is radioresistant; radiation therapy has no role in the routine management. Malignant fibrous histiocytoma is considered a part of the spectrum of osteosarcoma and is managed similarly.

CHONDROSARCOMA

Chondrosarcoma, which constitutes approximately 20 to 25% of all bone sarcomas, is a tumor of adulthood and old age with a peak incidence in the fourth to sixth decades of life. It has a predilection for the flat bones, especially the shoulder and pelvic girdles, but can also affect the diaphyseal portions of long bones. Chondrosarcomas can arise de novo or as a malignant transformation of an enchondroma or, rarely, of the cartilaginous cap of an osteochondroma. Chondrosarcomas have an indolent natural history and typically present as pain and swelling. Radiographically, the lesion may have a lobular appearance with mottled or punctate or annular calcification of the cartilaginous matrix. It is difficult to distinguish low-grade chondrosarcoma from benign lesions by x-ray or histologic examination. The diagnosis is therefore influenced by clinical history and physical examination. A new onset of pain, signs of inflammation, and progressive increase in the size of the mass suggest malignancy. The histologic classification is complex, but most tumors fall within the classic category. Like other bone sarcomas, high-grade chondrosarcomas spread to the lungs. Most chondrosarcomas are resistant to chemotherapy, and surgical resection of primary or recurrent tumors, including pulmonary metastases, is the mainstay of therapy. There are two histologic variants for which this rule does not hold, however. Dedifferentiated chondrosarcoma is a low-grade tumor that dedifferentiates into a high-grade osteosarcoma or a malignant fibrous histiocytoma, a tumor that responds to chemotherapy. Mesenchymal chondrosarcoma, a rare variant composed of a small cell element, also is responsive to systemic chemotherapy and is treated like Ewing's sarcoma.

EWING'S SARCOMA

Ewing's sarcoma, which constitutes approximately 10 to 15% of all bone sarcomas, is common in adolescence and has a peak incidence in the second decade of life. It typically involves the diaphyseal region of long bones and also has an affinity for flat bones. The plain radiograph may show a characteristic "onion peel" periosteal reaction with a generous soft tissue mass, which is better demonstrated byCT orMRI. This mass

is composed of sheets of monotonous, small, round, blue cells and can be confused with lymphoma, embryonal rhabdomyosarcoma, and small cell carcinoma. The presence of p30/32, the product of the *mic-2* gene (which maps to the pseudoautosomal region of the X and Y chromosomes) is a cell-surface marker for Ewing's sarcoma [and other members of a family of tumors called *peripheral primitive neuroectodermal tumors* (*PNETs*)]. Most PNETs arise in soft tissues; they include peripheral neuroepithelioma, Askin's tumor (chest wall), and esthesioneuroblastoma. Glycogen-filled cytoplasm detected by staining with periodic acid-Schiff is also characteristic of Ewing's sarcoma cells. The classic cytogenetic abnormality associated with this disease (and other PNETs) is a reciprocal translocation of the long arms of chromosomes 11 and 22, t(11;22), which creates a chimeric gene product of unknown function with components from the *fli-1* gene on chromosome 11 and *ews* on 22. This disease is very aggressive, and it is therefore considered a systemic disease. Common sites of metastases are lung, bones, and bone marrow. Systemic chemotherapy is the mainstay of therapy, often being used before surgery. Doxorubicin, cyclophosphamide or ifosfamide, etoposide, vincristine, and dactinomycin are active drugs. Local treatment for the primary tumor includes surgical resection, usually with limb salvage or radiation therapy. Patients with lesions below the elbow and below the mid-calf have a 5-year survival rate of 80% with effective treatment. Ewing's sarcoma is a curable tumor, even in the presence of obvious metastatic disease, especially in children less than 11 years old.

TUMORS METASTATIC TO BONE

Bone is a common site of metastasis for carcinomas of the prostate, breast, lung, kidney, bladder, and thyroid and for lymphomas and sarcomas. Prostate, breast, and lung primaries account for 80% of all bone metastases. Metastatic tumors of bone are more common than primary bone tumors. Tumors usually spread to bone hematogenously, but local invasion from soft tissue masses also occurs. In descending order of frequency, the sites most often involved are the vertebrae, proximal femur, pelvis, ribs, sternum, proximal humerus, and skull. Bone metastases may be asymptomatic or may produce pain, swelling, nerve root or spinal cord compression, pathologic fracture, or myelophthisis (replacement of the marrow). Symptoms of hypercalcemia may be noted in cases of bony destruction.

Pain is the most frequent symptom. It usually develops gradually over weeks, is usually localized, and often is more severe at night. When patients with back pain develop neurologic signs or symptoms, emergency evaluation for spinal cord compression is indicated (Chap. 102). Bone metastases exert a major adverse effect on quality of life in cancer patients.

Cancer in the bone may produce osteolysis, osteogenesis, or both. Osteolytic lesions result when the tumor produces substances that can directly elicit bone resorption (vitamin D-like steroids, prostaglandins, or parathyroid hormone-related peptide) or cytokines that can induce the formation of osteoclasts (interleukin 1 and tumor necrosis factor). Osteoblastic lesions result when the tumor produces cytokines that activate osteoblasts. In general, purely osteolytic lesions are best detected by plain radiography, but they may not be apparent until they are larger than 1 cm. These lesions are more commonly associated with hypercalcemia and with the excretion of hydroxyproline-containing peptides indicative of matrix destruction. When osteoblastic

activity is prominent, the lesions may be readily detected using radionuclide bone scanning (which is sensitive to new bone formation), and the radiographic appearance may show increased bone density or sclerosis. Osteoblastic lesions are associated with higher serum levels of alkaline phosphatase, and, if extensive, may produce hypocalcemia. Although some tumors may produce mainly osteolytic lesions (e.g., kidney cancer) and others mainly osteoblastic lesions (e.g., prostate cancer), most metastatic lesions produce both types of lesion and may go through stages where one or the other predominates.

In older patients, particularly women, it may be necessary to distinguish metastatic disease of the spine from osteoporosis. In osteoporosis, the cortical bone may be preserved, whereas cortical bone destruction is usually noted with metastatic cancer.

Treatment of metastatic bone disease depends on the underlying malignancy and the symptoms. Some metastatic bone tumors are curable (lymphoma, Hodgkin's disease), and others are treated with palliative intent. Pain may be relieved by local radiation therapy. Hormonally responsive tumors are responsive to hormone inhibition (antiandrogens for prostate cancer, antiestrogens for breast cancer). Strontium 89 and samarium 153 are bone-seeking radionuclides that can exert antitumor effects and relieve symptoms. Bisphosphonates such as pamidronate may relieve pain and inhibit bone resorption. Monthly administration prevents bone-related clinical events and may reduce the incidence of bone metastases in women with breast cancer. When the integrity of a weight-bearing bone is threatened by an expanding metastatic lesion that is refractory to radiation therapy, prophylactic internal fixation is indicated. Overall survival is related to the prognosis of the underlying tumor. Bone pain at the end of life is particularly common; an adequate pain relief regimen including sufficient amounts of narcotic analgesics is required.**The management of hypercalcemia is discussed in Chap. 341.*

(Bibliography omitted in Palm version)

Back to Table of Contents

99. METASTATIC CANCER OF UNKNOWN PRIMARY SITE - *Richard M. Stone*

INCIDENCE AND EPIDEMIOLOGY

The presenting findings in a patient with a newly discovered malignancy may not reveal its site of origin. Patients with cancer of unknown primary site (CUPS) present difficult diagnostic and therapeutic dilemmas. First, as additional studies may be many, costly, and/or uncomfortable for the patient, the strategy used in searching for the primary must assess what, if any, result the identification of the site of origin would have on the patient's treatment and survival. Second, while individuals with CUPS fare poorly overall (median survival is 4 to 11 months), certain subgroups of patients are more likely to benefit from treatment and, in some cases, to enjoy long disease-free survival. The literature is a poor guide for the care of such patients, owing to the heterogeneity of the tumors, the selection bias in small retrospective studies, and the variability in the definition of the syndrome and the thoroughness of the evaluation performed to identify a primary site.

No universally accepted definition of theCUPSsyndrome exists. An occult neoplasm should fulfill all of the following criteria: (1) biopsy-proven malignancy; (2) unrevealing history, physical examination, chest film, abdominal and pelvic computed tomography (CT) scans, complete blood counts, chemistry survey, mammography (women),b human chorionic gonadotropin (bhCG) levels (men), afetoprotein (AFP) levels (men), and prostate-specific antigen (PSA) levels (men); (3) histologic evaluation not consistent with a primary tumor at the biopsy site; and (4) failure of additional diagnostic studies (based only on findings from the laboratory and pathologic review) to identify the primary site. Such additional diagnostic tests could include, for example, colonoscopy in a patient whose rectal examination discloses guaiac-positive stool or a meticulous otolaryngologic examination in a patient who presents with squamous cell carcinoma in a cervical node. Many cases that fulfill the definition of CUPS offer clues that a given organ is the probable site of origin. Epidemiologic data suggest that the incidence of cancers for which the primary site is unknown is decreasing. CUPS accounts for about 2% of all cancer diagnoses -- about 24,400 cases in the year 2000. Most patients with CUPS are over age 60.

BIOLOGIC CONSIDERATIONS

The biologic behavior ofCUPS is unique. In ~25% of patients, the primary site becomes apparent during the course of the illness; in about 57% of patients, the primary site can be diagnosed at autopsy; but in almost 20%, the primary site remains obscure even at autopsy. Cancers presenting as CUPS often display unusual patterns of metastatic spread (e.g., pancreatic cancer presenting with bony metastases). The fact that more tumor bulk is present at distant sites than in the tissue of origin suggests that the genetic lesions underlying cases of CUPS produce a distinctly aggressive phenotype. Microsattelite DNA analysis has shown that the same pattern of genetic alterations that appear in a cervical lymph node metastasis can be found in seemingly morphologically normal aerodigestive tissue. Such data imply that clinically evident metastases may be able to arise from microscopic primary lesions. Although physiologic and genetic data that might account for the distinctive natural history of CUPS neoplasms are scant, cell lines derived from such tumors may have abnormalities of chromosome 1, a finding

generally associated with advanced malignancy. In some patients, the primary tumor spontaneously regresses (perhaps under immunologic attack) or necroses. In some, a primary lesion was resected years before presentation (e.g., melanoma).

CLINICAL PRESENTATION, DIAGNOSTIC EVALUATION, AND PATHOLOGY

History and Physical Examination Patients present with a variety of symptoms and signs, including fatigue, weight loss, other systemic symptoms, pain, abnormal bleeding, abdominal swelling, subcutaneous masses, and lymphadenopathy. OnceCUPS is considered, the physician's approach must involve reasonable efforts to identify the primary site or to determine the histology or subcategory of the metastatic tumor to decide on the optimal therapy. Though usually unrevealing, a thorough history and physical examination should be carried out to elicit easily obtainable clues regarding the primary site. The patient should be questioned concerning epigastric pain, which, if present, would mandate careful exclusion of pancreatic carcinoma as well as other gastrointestinal malignancies. Symptoms referable to a given location (e.g., new cough, hematochezia, hemoptysis, change in bowel habits, unusual vaginal bleeding, nipple discharge) should prompt an aggressive specific diagnostic approach. Occupational exposure to asbestos, for example, would raise the suspicion of mesothelioma. The absence of prior smoking reduces the likelihood of lung cancer but does not exclude it. A history of fulguration of a skin lesion, colonic polypectomy, dilatation and curettage, or prostate biopsy should prompt a review of the original histology.

Pathology Review The most important aspect of the workup of a patient withCUPS is the thorough evaluation of the tissue obtained at biopsy by light-microscopy, immunohistochemistry, ultrastructural studies, immunophenotyping, and karyotypic and molecular biologic analysis. First, if the original biopsy sample is inadequate for either confirmation of malignancy or the performance of additional specialized studies, rebiopsy is mandatory. The clinician must have a close working relationship with a pathologist skilled in the evaluation of tumor specimens, especially when the organ of origin is uncertain. Plans may be made to process the tissue for (1) routine light-microscopy, histochemical, and immunohistochemical analysis; (2) freezing for DNA and RNA isolation or for in situ genetic and immunologic evaluation; and (3) special fixation for ultrastructural analysis. Single-cell tumor suspensions in short-term culture permit cytogenetic analysis.

If routine histologic analysis fails to suggest the tissue of origin (e.g., gland formation in adenocarcinoma, psammoma bodies in ovarian or thyroid cancer, or spindle architecture in sarcomas), special histochemical studies may be helpful. For example, mucin positivity is helpful in recognizing a poorly differentiated adenocarcinoma. Light-microscopic analysis will show approximately 60% ofCUPStumors to be well or moderately differentiated adenocarcinomas, 30% poorly differentiated carcinomas/adenocarcinomas, and 5% poorly differentiated malignant neoplasms not further classifiable. In the poorly differentiated neoplasms, immunohistochemical, cytogenetic, and molecular biologic studies can be extremely useful in identifying sarcomas, germ cell carcinomas, lymphomas, neuroendocrine neoplasms (including melanoma), and other tumors whose diagnosis would suggest a more specific therapeutic approach.

Immunohistochemical Analysis Antibodies to specific cell components make it possible to characterize tumors that are not identified by standard techniques.Table 99-1 provides a list of antigens that may be assessed in undifferentiated or poorly differentiated specimens. A diagnosis of lymphoma should be excluded by employing antibodies reactive to the leukocyte common antigen (LCA, CD45). LCA-positive tumors are lymphomas and have the same chances of responding to therapy as if the diagnosis were unambiguous. About half of patients with aggressive-histology lymphoma can be cured with combination chemotherapy (Chap. 112). The immunohistochemical detection of specific types of filament proteins is helpful in the identification of carcinomas and sarcomas. The presence of keratin suggests carcinoma; all epithelial tumors contain this protein. Specific types of cytokeratins (CK) permit a firm diagnosis. For example, ovarian cancers are CK20-/CK7+, colorectal cancers are CK20+/CK7-, and pancreaticobiliary tumors are CK20+/CK7+. However, certain sarcomas, mesotheliomas, and germ cell tumors are also keratin-positive. Sarcomas may react with antibodies to desmin. Sarcoma subgroups may be identified by expression of myoglobin (rhabdomyosarcoma) or factor VIII (angiosarcoma or Kaposi's sarcoma). Prostate, breast, and thyroid carcinomas express, respectively,PSA, gross cystic fluid protein, or thyroglobulin. The finding ofAFP,bhCG, or placental alkaline phosphatase staining is very helpful in assigning a germ cell origin. The S-100 protein is present in virtually all primary and metastatic melanomas, including the amelanotic variety. However, S-100 positivity is also found in other tumors of neuroendocrine origin (e.g., small cell lung cancer, carcinoid, neuroepithelioma); a more specific marker for melanomas is the HMB45 (human melanoma black) antigen.

Other Diagnostic Approaches Electron microscopy can identify cell junctions (i.e., desmosomes, typical of epithelial cancers), neuroendocrine granules, melanosomes, and muscle filaments. Cytogenetic analysis may identify tumors with specific chromosomal translocations or other genetic abnormalities (Table 99-1). Cytogenetic abnormalities can also be determined by fluorescence in situ hybridization with chromosome-specific probes, a technique that does not require cells to divide, as is the case with traditional karyotype analysis. Fresh tissue may be required for detection of estrogen or progesterone receptors (to assess breast cancer) or antigens that are sensitive to fixation. Lineage can be assigned by analysis of DNA for signature gene rearrangements, such as those of immunoglobulin (B cell) or T cell receptor (T cell). Technological advances promise to influence the diagnosis of cancer. Isolation of mRNA from tumor specimens may permit the molecular profiling of tumors by microarray analysis of gene expression. This could lead to novel classifications of tumors based on molecular characteristics that may predict clinical behavior and/or response to specific therapies.

Additional Studies If the pathologist does not identify the likely tissue of origin, it is unlikely that additional expensive diagnostic tests will benefit the patient. In females with metastatic adenocarcinoma or poorly differentiated carcinoma, mammography should be performed, although the diagnostic yield will be quite low except in patients with axillary metastases. Magnetic resonance imaging, positron-emission tomography, or indium-111-pentreotide scanning can identify occult primary breast lesions but are expensive. The use of abdominal/pelvicCT scans leads to the identification of the primary site (often the pancreas) in up to 35% of patients but has little effect on natural history. Whether to measure serum tumor markers such asAFP,bhCG,

carcinoembryonic antigen (CEA), CA-125 (associated with ovarian cancer), andPSA is controversial; value has not been proven. Numerous studies have shown a lack of benefit of contrast studies (upper gastrointestinal series, barium enema, or intravenous pyelogram) in patients withCUPS who have no specific symptoms and no findings referable to the gastrointestinal or urinary tract. Moreover, autopsy series reveal that the most likely primary site of origin includes epithelial tissues such as lung, stomach, colon, and kidney, which give rise to tumors that respond poorly to chemotherapy, minimizing the therapeutic impact of such a diagnosis.

Additional invasive diagnostic studies are indicated if the presentation strongly suggests a particular primary site. For example, radiographic evidence of lung or mediastinal involvement would mandate fiberoptic bronchoscopy to exclude lung cancer. In the relatively unusual case of metastatic squamous cell cancer presenting in an inguinal lymph node, anoscopy and colposcopy should be performed to detect carcinoma of the vulva, cervix, vagina, penis, or anus, all of which may be cured even with lymph node spread. A summary of a reasonable diagnostic approach is found inTable 99-2.

TREATMENT

Prognostic Subgroups The exclusion of treatable and potentially curable neoplasms is important. Patients with squamous cell carcinoma have a somewhat longer median survival (9 months) than do those with adenocarcinoma or unclassifiable neoplasms (4 to 6 months). If laboratory studies indicate a significant likelihood that the neoplasm is a lymphoma, germ cell tumor, sarcoma, neuroendocrine tumor, or breast or prostate cancer, then disease-appropriate therapy should be administered. Patients with lymphoma or a germ cell neoplasm may be cured with combination chemotherapy. In other malignancies, effective palliative chemotherapy (for sarcoma or a breast or neuroendocrine tumor) or hormonal therapy (for breast or prostate cancer) should be strongly considered. Although often requiring electron microscopy for diagnosis, neuroendocrine tumors (especially if anaplastic) often respond to cisplatin-based chemotherapy.

Patients in whom the primary site can be identified fare somewhat better than those in whom it remains undefined. Classification and regression tree (CART) analysis has led to a prognostic index ranging from a median survival of 40 months (those with one or two organ sites involved; not adenocarcinoma in histology; and without adrenal, bone, liver, or pleural involvement) to a median survival of 5 months (liver metastases, nonneuroendocrine histology, age ³62). Patients may often be categorized as having one of several clinical features or syndromes suggesting a specific form of potentially beneficial therapy (Table 99-3).

Syndrome of Unrecognized Extragonadal Germ Cell Cancer A subset of patients with poorly differentiatedCUPS are responsive to chemotherapy. These patients display one or more of the following features: age <50; tumor involving midline structures, lung parenchyma, or lymph nodes; an elevated serumAFP orbhCGlevel; evidence of rapid tumor growth; or tumor responsiveness to previously administered radiotherapy or chemotherapy. Platinum-based chemotherapy has led to long-term survival in a fraction of patients with these features, especially those who have a favorable performance status at diagnosis, suggesting that their tumors behaved like germ cell neoplasms. If all

patients with poorly differentiated carcinoma (including poorly differentiated adenocarcinoma) are treated with a chemotherapy regimen designed for germ cell cancer (e.g., cisplatin plus etoposide or vinblastine, often also with bleomycin) (Chap. 96), about 25% will respond completely and 33% will have a partial response. Patients whose disease does not respond to two cycles of therapy should not continue therapy. One in six patients survives>5 years without evidence of disease. Patients with poorly differentiated carcinoma or adenocarcinoma whose tumors have abnormalities of chromosome 12 similar to those described in patients with proven germ cell cancer are more likely to respond to platinum-based chemotherapy than are patients with a similar presentation whose tumors lack this cytogenetic abnormality.

Peritoneal Carcinomatosis in Women Women who present with increased abdominal girth and a pelvic mass or pain and who are found to have adenocarcinoma throughout the peritoneal cavity without a clear site of origin also may benefit from platinum-based chemotherapy. This syndrome has been termed *primary peritoneal papillary serous carcinoma* or *multifocal extraovarian serous carcinoma*. While breast cancer or a gastrointestinal malignancy can produce these findings, peritoneal carcinomatosis is most commonly ascribed to ovarian cancer, even in patients with apparently normal ovaries at the time of laparotomy. Especially if psammoma bodies or a papillary configuration is noted in the pathology examination or if the CA-125 level is elevated, women with adenocarcinoma of the peritoneal cavity without a defined primary should receive maximum surgical cytoreduction followed by cisplatin (or carboplatin) plus paclitaxel. The stage-specific response to such therapy appears to be comparable to that for patients with proven ovarian cancer. About 10% of patients who present in this fashion may remain free of disease 2 years after diagnosis.

Carcinoma in an Axillary Lymph Node in a Female Women with adenocarcinoma or poorly differentiated carcinoma in an axillary mass should receive treatment for stage II breast cancer whether or not a careful breast examination or mammography suggests the diagnosis of primary breast cancer and whether or not estrogen or progesterone receptors are detectable in the node. Even if no lesion is found in the breast, a breast recurrence will develop in one-half of these patients if no mastectomy is performed. Modified radical mastectomy or breast irradiation reduces the risk of local recurrence. In addition, adjuvant systemic therapy (chemotherapy and/or tamoxifen, depending on menopausal and estrogen receptor status) should be given to reduce the risk of developing evident metastatic breast cancer (Chap. 89). Adjuvant systemic therapy may be administered before definitive local radiation treatment. Women with axillary metastases without an obvious breast primary appear to have the same likelihood of prolonged disease-free survival as patients with typical stage II breast cancer.

Bone Metastases in Males Particularly if the lesions are osteoblastic, the serumPSAlevel should be measured, as the probability of prostate carcinoma is high. Empirical hormonal therapy (e.g., leuprolide and flutamide) should be strongly considered.

Cervical Lymph Node Metastases Patients who present with a neck mass should be considered to have a primary tumor of the upper aerodigestive tract (*head and neck cancer*) until a different source is proven. Especially if the pathologist diagnoses squamous histology and the node is located in a high or midcervical area, a careful ear,

nose, and throat examination including direct laryngoscopy, nasopharyngoscopy, and random blind biopsies should be undertaken. A thyroid examination and scan should be performed to rule out a primary thyroid tumor, especially if the histology is not definitely squamous. Definitive local therapy (external beam radiation or radical neck dissection) combined with platinum-based chemotherapy may lead to prolonged survival in those with head and neck primaries (Chap. 87).

Adenocarcinoma and Liver Metastases Liver metastases from an adenocarcinoma is not as well characterized as a syndrome as the unrecognized germ cell cancer syndrome (nor as responsive to therapy). However, such patients may have a primary stomach, biliary, or colorectal tumor. Tumors with limited hepatic involvement may be amenable to resection. A flexible sigmoidoscopy or colonoscopy may detect a potentially obstructive colonic lesion. If a tumor is found, resection may be beneficial, depending on the tumor's size; even if none is found, treatment with a combination of 5-fluorouracil plus leucovorin is palliative for some patients with presumed metastatic gastrointestinal malignancy. Given the severe diarrhea that may be a consequence of this regimen and the relative resistance of gastrointestinal tumors to chemotherapy, patients should be informed of the risks before treatment.

Patients not falling into one of the preceding categories should be treated palliatively. In some patients, observation is appropriate. For example, individuals without evidence of additional metastatic disease who have undergone resection of a solitary pulmonary nodule containing malignant cells may actually have undergone definitive therapy for a small primary lung tumor. Radiation therapy may relieve symptoms in patients with bony pain or neurologic compromise. The largest and most poorly responsive subgroup are those with moderate to well-differentiated adenocarcinomas. Combination chemotherapy is frequently employed in such patients; however, response rates to "all-purpose" regimens [e.g., FAM (5-fluorouracil, doxorubicin, mitomycin C), FACP (5-fluorouracil, doxorubicin, cyclophosphamide, cisplatin)], or to ICE (ifosfamide, carboplatin, etoposide) are generally well under 50%, especially if patients with poorly differentiated adenocarcinoma, who have a higher response rate, are excluded; complete responses are rare. Regimens containing mitomycin C are associated with the risk of hemolytic uremic syndrome. In some series, patients with a good performance status whose disease is limited to soft tissue sites or extends only above the diaphragm have shown a better rate of response to therapy. While patients whose disease responds to treatment seem to have better survival than those whose disease does not respond, the difference may be related to inherent characteristics of the tumor rather than to a beneficial effect of chemotherapy.

Before combination chemotherapy is attempted in a patient withCUPS, the potential benefits must be weighed carefully against the certainty of toxicity. While some randomized studies have reported a benefit of one form of therapy over another, these reports are generally plagued by small numbers of patients and inadequate control of potential prognostic variables. Depending on motivation, eligibility, and availability, patients with CUPS may be candidates for evaluation of new (phase I) therapies.

(Bibliography omitted in Palm version)

Back to Table of Contents

100. PARANEOPLASTIC SYNDROMES - *Bruce E. Johnson*

ENDOCRINE SYNDROMES

Paraneoplastic syndromes are caused by factors produced by cancer cells that often act at a distance from both the primary cancer site and its metastases. The accurate documentation of a paraneoplastic endocrine syndrome requires (1) demonstration of mRNA expression and protein production by the tumor tissue, (2) biochemical and clinical resolution of the syndrome following successful surgical resection, (3) elevated levels of the hormone in the peripheral blood, (4) a tenfold or greater concentration gradient in the blood before and after it passes through the cancer, and (5) normal or suppressed endogenous hormone production.

The three major classes of hormones are steroids, monoamines, and peptides/proteins. Production of steroid hormones by malignant tumors is rare; lymphomas may produce 1,25-dihydroxyvitamin D from circulating 1-hydroxyvitamin D, but most other steroid-producing tumors are benign tumors of the glands that normally secrete the steroid. Monoamines such as norepinephrine and epinephrine are secreted by pheochromocytomas (Chap. 332), but they are not secreted ectopically by malignant tumor cells.

Most hormonal syndromes in patients with cancer are related to the production of peptide or protein hormones. The most common of these endocrine syndromes are listed inTable 100-1, together with the protein hormones that mediate them and the tumors that most commonly produce the hormones. A peptide hormone generally is encoded by an mRNA that is translated into a larger prohormone molecule, which undergoes a number of posttranslational modifications, including cleavage, glycosylation, and/or other steps. For example, pro-opiomelanocortin can be cleaved to yield adrenocorticotropic hormone (ACTH), lipotropin, endorphin, melanocyte-stimulating hormone, and/or enkephalin, with different cell types producing different products. In addition, some cells use alternatively spliced forms of the message to produce different proteins (e.g., calcitonin vs. calcitonin gene-related peptide).

Tumor cells of nonendocrine organs often lack certain components of the pathway that leads from prohormone to biologically active hormone to secreted product. Generally, as a result of defects in protein processing or post-translational changes, tumor cells may produce proteins that are structurally related to but biologically less active than the normal hormones. Thus, cancer patients may have elevated levels of immunoreactive hormones in plasma in the absence of clinical syndromes of hormone excess.

The severity of paraneoplastic endocrine syndromes often parallels the clinical course of the cancer. However, with some benign or slowly growing tumors, the hormone syndrome can be the major cause of morbidity. Despite their production by tumor cells, hormones are not very good tumor markers. Human chorionic gonadotropin (hCG) is a reliable tumor marker in some forms of testicular cancer, but no other hormone is used to quantitate tumor mass.

Most endocrine cancer syndromes occur with tumors derived from neuroendocrine or neural crest tissue (small cell lung cancer, carcinoid tumors). The genetic mechanisms

that account for the production of a hormone by a cell that does not usually produce it are not clear. Oncogenes may activate other cellular genes (including genes that encode hormones) that normally are silent. Alternatively, demethylation of normally methylated inactive genes may permit expression in rapidly dividing cells.

HYPERCALCEMIA OF MALIGNANCY

Hypercalcemia of malignancy, the most common paraneoplastic endocrine syndrome, is responsible for approximately 40% of all hypercalcemia (Chap. 341). Hypercalcemia with cancer is classified as humoral hypercalcemia of malignancy (HHM), which is caused by circulating hormones, or local osteolytic hypercalcemia (LOH), which is caused by local paracrine factors secreted by cancers within bone. Parathyroid hormone-related peptide (PTHrP) causes nearly all cases of HHM, while the mediators of LOH in bone are heterogeneous.

Pathogenesis Eighty percent of patients with hypercalcemia of malignancy haveHHM.PTHrP is composed of 139 to 173 amino acids; 8 of the first 13 amino acids at the amino-terminal end are identical to the amino-terminal portion of parathyroid hormone (PTH). PTHrP binds to PTH receptors in the bone and kidney and causes increased bone resorption, decreased bone formation, increased renal tubular reabsorption of calcium, increased phosphaturia, and increased urinary levels of cyclic adenosine monophosphate, leading to hypercalcemia. PTHrP is detected in the plasma in ~80% of cancer patients with hypercalcemia. Rare patients have been reported to have hypercalcemia caused by ectopically produced authentic PTH. HHM in lymphoma may be caused by the production of 1,25-dihydroxyvitamin D by the tumor.

Twenty percent of patients with hypercalcemia haveLOH, in which hypercalcemia is caused by the local production of hormones or cytokines by cancers that have spread to the bone or bone marrow; such factors increase bone resorption in the area around the cancer. The ectopically produced hormones that may play a role in LOH include transforming growth factorsa and b, interleukin (IL)-1, IL-6, prostaglandins, and tumor necrosis factor.

Clinical Manifestations The initial symptoms and signs of hypercalcemia (calcium level[3] 2.6 mmol/L)

[1]Calcium measurements given in millimoles per liter can be multiplied by 4 to convert to milligrams per deciliter or by 2 to convert to milliequivalents per liter. include malaise, fatigue, confusion, anorexia, bone pain, polyuria, polydipsia, weakness, constipation, nausea, and vomiting. Neurologic symptoms and signs in profound hypercalcemia (>3.5 mmol/L) include confusion, lethargy, coma, and death. The cancers associated withHHM are non-small cell lung cancer and cancers of the breast, kidney, head and neck, and bladder. HHM is particularly common in patients with cancers of squamous cell histology. Hypercalcemia is uncommon at presentation (<1% of patients) but becomes more common as the cancer progresses and is present in 10 to 20% of patients near the time of death.LOH is responsible for hypercalcemia in patients with breast cancer, myeloma, lymphoma, and leukemia. Among hypercalcemic patients with breast cancer, approximately half have HHM and half have LOH.

Diagnosis The patient with cancer who develops hypercalcemia should be evaluated for other causes of hypercalcemia, including use of thiazide diuretics, vitamin D, or lithium, hyperthyroidism, and sarcoidosis. If the underlying cancer is controlled, elevation of serumPTH as measured by immunoassay suggests primary hyperparathyroidism, which may be responsible for as many as 10% of cases of hypercalcemia of cancer and which should be treated like other cases of hyperparathyroidism (Chap. 341). A normal PTH level and a low serum phosphorus level in the absence of bone metastases support the diagnosis ofHHM, while a normalPTHrPlevel and normal phosphorus in a patient with bone metastases suggestLOH.

TREATMENT

The median survival of patients with hypercalcemia of malignancy is only 1 to 3 months. Intervention to reverse hypercalcemia should be undertaken when the cancer is likely to be controlled with appropriate systemic or local treatment.

The treatment ofHHM andLOH is similar (Chap. 341). Patients with mild to moderate hypercalcemia (2.7 to 3.5 mmol/L) can be treated with 2 to 4 L of saline hydration per day and furosemide to prevent intravascular volume overload. The bisphosphonate pamidronate (90 mg intravenously) decreases osteoclastic bone resorption. Combined administration of diuretics and pamidronate reduces the serum calcium to normal values in 90% of patients within 7 days. Doses may be repeated as needed. In patients with LOH, glucocorticoids may inhibit the production of cytokines that promote bone resorption. Severe hypercalcemia [>3.50 mmol/L (>14 mg/dL)] with alteration of mental status can be treated with all of the above plus salmon calcitonin, 4 to 8 U/kg, administered intramuscularly or subcutaneously every 12 h. Calcitonin administration will decrease the serum calcium within 24 h, and its hypocalcemic effect can be prolonged in patients with LOH by adding glucocorticoids. If these agents are not effective in reducing the serum calcium, plicamycin and gallium nitrate may be added.

HYPONATREMIA OF MALIGNANCY

Hyponatremia of malignancy (Na^+ level<130 mmol/L) is usually due to the inappropriate secretion of arginine vasopressin (AVP) and is termed the *syndrome of inappropriate antidiuretic hormone secretion* (SIADH). In rare cases, atrial natriuretic peptide produces hyponatremia.

Pathogenesis Small cell lung cancer is the malignancy chiefly responsible for producing ectopicAVP. AVP mRNA is expressed and translated, and the product is processed into the nonapeptide AVP, which is secreted into the circulation. The ectopically produced AVP binds to receptors in the kidney, causing retention of free water with resulting hypoosmolality in the plasma and hyperosmolality in urine (Chap. 329).

About 15% of cancer patients withSIADH do not have evidence of ectopic production ofAVP. In some of these patients, tumors secrete atrial natriuretic peptide. This hormone inhibits sodium reabsorption in the proximal tubule and inhibits release of renin and aldosterone. How atrial natriuretic peptide leads to hyponatremia is not clear.

Clinical Manifestations SIADH is commonly recognized as asymptomatic hyponatremia on routine serum chemistry examination. It is present at the time of diagnosis in 15% of patients with small cell lung cancer, 3% of patients with head and neck cancer, and<1% of patients with non-small cell lung cancer. Hyponatremia may also occur with primary brain tumors, hematologic malignancies, melanoma, sarcoma, and gynecologic, gastrointestinal, breast, prostate, and bladder cancers. The symptoms of mild hyponatremia (>120 mmol/L) include difficulty focusing attention, fatigue, nausea, vomiting, anorexia, weakness, and headache. Profound hyponatremia (<120 mmol/L) can cause confusion, lethargy, coma, seizures, and death.

Diagnosis (See also Chap. 329) SIADH is suspected in patients with hyponatremia (serum sodium<130 mmol/L) and a concentrated urine (osmolality>300 mosm/kg). Patients are euvolemic, are not using diuretics, and have normal thyroid and adrenal function. Polydipsia is excluded by the urine osmolality. Pseudohyponatremia can be present if serum glucose, triglyceride, or protein levels are high. Conditions other than cancer that can cause SIADH include central nervous system disorders, pulmonary infections, positive-pressure breathing, pneumothorax, asthma, and a wide array of drugs, including chemotherapeutic agents (vincristine, vinblastine, cisplatin, cyclophosphamide, melphalan), thiazide diuretics, carbamazepine, antidepressants, nicotine, and narcotics.

TREATMENT

Treatment should be directed at the underlying cancer. Patients whose tumors have not been or cannot be controlled are candidates for restriction of fluid intake to 500 mL/d. Such restriction corrects hyponatremia within 7 days in most patients, but it is difficult and uncomfortable for patients to maintain fluid restriction for extended periods. Oral demeclocycline (600 to 1200 mg/d) may be useful in blocking the effects of AVP but can cause renal insufficiency. Other agents that may be used for the treatment of hyponatremia include dilantin and lithium.

Rare patients develop profound hyponatremia and altered mental status. These patients should be treated with normal saline hydration and furosemide diuresis. If that is not effective, 3% saline can be administered via a central line together with furosemide diuresis to prevent hypervolemia. Hypertonic saline is rarely required and must be given slowly; fluid balance and electrolytes should be monitored several times per day, and the increase in sodium should be limited to 0.5 mmol/L per hour to prevent pontine lysis (Chap. 329).

ECTOPIC ACTH SYNDROME

Ectopic production of ACTH by cancer cells is responsible for ~15% of all cases of Cushing's syndrome and for most cases of Cushing's syndrome that occur in cancer patients (Chaps. 328,331). In rare patients, Cushing's syndrome is caused by ectopically produced corticotropin-releasing hormone (CRH), which stimulates pituitary ACTH release.

Pathogenesis When pro-opiomelanocortin mRNA is expressed in cancer cells, the

241-amino-acid prohormone is translated and processed into a variety of molecules, including, in some cases, the 39-amino-acid hormoneACTH, which can be secreted into the circulation. The ectopically produced ACTH causes excessive secretion of glucocorticoids and mineralocorticoids by the adrenals.

Clinical Manifestations Women make up 50% of patients with ectopicACTHsyndrome and 90% of patients with pituitary Cushing's disease. Therefore, Cushing's syndrome in men is more likely to be caused by ectopic ACTH than by a pituitary tumor. Because of mineralocorticoid excess, patients with ectopic Cushing's syndrome usually have hypokalemic alkalosis at presentation, a rare finding in patients with Cushing's disease. Other common manifestations of ectopic ACTH syndrome include weakness, hypertension, and hyperglycemia. Ectopic ACTH syndrome in patients with slow-growing cancers (e.g., carcinoids) may develop typical features of central obesity, moon facies, hyperpigmentation, and hirsutism in addition to the metabolic abnormalities.

EctopicACTHsyndrome is most commonly due to small cell lung cancer (50% of cases), bronchial carcinoid tumors (10%), thymic carcinoid tumors or thymomas (10%), pancreatic islet cell tumors (10%), pheochromocytoma or other neural crest tumors (5%), or medullary carcinoma of the thyroid (5%). About 2% of patients with small cell lung cancer and bronchial carcinoids have ectopic ACTH syndrome at the time of diagnosis. Patients with small cell lung cancer and ectopic ACTH syndrome have shorter survival rates than patients without the syndrome and are more likely to develop opportunistic infections.

Diagnosis (See alsoChaps. 328,331) EctopicACTHsyndrome is usually characterized by elevated levels of urinary free cortisol that do not decrease after administration of high doses of dexamethasone (8 mg/d). However, in 20 to 30% of patients with ectopic ACTH syndrome, urinary cortisol levels decrease by >50% after administration of high-dose dexamethasone. The plasma levels of ACTH are markedly elevated in more than half of patients. If these tests do not provide definitive evidence of ectopic ACTH syndrome, bilateral inferior petrosal vein sampling will show an elevated ACTH level in petrosal vein blood that does not increase after administration ofCRH.

TREATMENT

Treatment of the ectopicACTHsyndrome should be directed at the underlying cancer: chemotherapy for small cell lung cancer; surgical resection or radiation therapy for carcinoids. Some patients with ectopic Cushing's syndrome have no evidence of tumor after extensive evaluation. These patients should be treated symptomatically and followed closely with periodic imaging studies, because they may have slow-growing tumors amenable to surgical resection.

Agents that inhibit steroidogenesis in the adrenal gland include ketoconazole (400 to 1200 mg/d), which reduces urinary cortisol excretion by more than half in two-thirds of patients, and metyrapone (1 to 4 g/d), which also reduces urinary cortisol excretion. Patients who are in good condition and whose manifestations are not controlled by drugs may be considered for adrenalectomy.

ECTOPIC ACROMEGALY

Ectopic production of growth hormone-releasing hormone (GHRH) is the predominant cause of ectopic acromegaly (Chap. 328).

PathogenesisGHRH is processed into 40- and 44-amino-acid peptides and binds to receptors in the pituitary, increasing production of growth hormone that increases insulin-like growth factor (IGF)-1 production in peripheral tissues. Rare cases of ectopic acromegaly are due to ectopic production of growth hormone itself by tumors.

Clinical Manifestations The symptoms and signs of ectopic acromegaly develop over several years and include increasing glove and shoe size, facial disfigurement, arthralgias, amenorrhea-galactorrhea or impotence, hypertension, muscle weakness, and diabetes mellitus. Ectopic acromegaly has been reported in fewer than 100 patients and accounts for 1% or less of all cases of acromegaly. The cancers associated with ectopic acromegaly include carcinoid tumors of the bronchus, pancreatic islet cell tumors, and cancers of the lung, breast, colon, and adrenal glands.

Diagnosis If a clinical diagnosis of acromegaly is suspected in a patient with cancer, the serum levels of GHRHandIGF-1 and the glucose-suppressed growth hormone (GH) serum level should be measured (Chap. 328). Patients with elevated GHRH levels and acromegaly have ectopic acromegaly. Patients without evidence of cancer who have elevated GHRH levels should undergo imaging of the central nervous system, chest, and abdomen to look for an occult cancer. Patients with cancer, low GHRH levels, high GH levels, and elevated IGF-1 levels should undergo magnetic resonance imaging of the pituitary and hypothalamus. If no pituitary tumor is detected, GH may be secreted directly by the known tumor. Not all GH -secreting tumors of the pituitary are demonstrable by imaging techniques, however.

TREATMENT

The therapy of ectopic acromegaly should be directed at the underlying cancer and should consist of surgical resection or radiation therapy for patients with carcinoid and islet cell tumors. Medical control of ectopic acromegaly is obtained by using octreotide (100 to 250 ug every 8 h), which inhibits pituitary secretion of growth hormone. Octreotide produces symptomatic improvement in approximately two-thirds of patients.

GYNECOMASTIA

Ectopic production ofhCG or estrogens by tumors such as cancers of the lung and testis is responsible for approximately 3% of cases of gynecomastia detected in men (Chap. 337). Ectopic production of hCG is the most common cause of paraneoplastic gynecomastia; the hCG acts by stimulating the Leydig cells of the testis to produce increased amounts of estrogen. Alternatively, on rare occasions, a tumor (such as a hepatoma or a germ cell tumor with choriocarcinoma elements) contains aromatase enzyme activity that converts circulating androgens to estrogen. Leydig cell or Sertoli cell tumors may also secrete estradiol. In all cases, the increased ratio of estrogen to testosterone leads to the proliferation of breast tissue and gynecomastia. Other tumors rarely associated with ectopic gynecomastia include carcinoid tumors of the bronchus,

intestine, and small cell lung cancer.

About 5% of men with testicular choriocarcinoma present with an enlarging breast mass. In the absence of an obvious cancer, men presenting with gynecomastia should have a careful examination of the testes and measurement of serumhCG. Patients with a testicular mass should undergo an inguinal orchiectomy for diagnosis and treatment. If no testicular mass is found by physical examination, the testes should be examined with ultrasound. Patients with an elevated hCG level and no testicular mass should undergo evaluation for an extragonadal germ cell tumor.

TREATMENT

The therapy of tumor-associated gynecomastia should be directed at the underlying cancer: chemotherapy is used for testicular cancers, and surgical resection or radiation therapy for carcinoids and islet cell tumors. In patients with successfully treated testicular cancer, gynecomastia completely resolves in three-fourths of cases.

NON-ISLET CELL TUMOR HYPOGLYCEMIA

Hypoglycemia that is not caused by the ectopic production of insulin (as in patients with islet cell tumors of the pancreas) can occur with large, slow-growing sarcomas, mesotheliomas, and hepatomas (Chap. 334). Ectopic production ofIGF-2 is responsible for hypoglycemia in most patients with non-islet cell tumors. The ectopically produced IGF-2 inhibits glycogenolysis and gluconeogenesis in the liver, suppresses lipolysis, and increases peripheral glucose utilization, thereby causing hypoglycemia. IGF-2 may also act as an autocrine growth factor for the tumor.

Patients with large sarcomas (1 to 10 kg) may develop hypoglycemia, particularly with fasting. Headache, fatigue, confusion, or seizures may occur. Patients with a large sarcoma and hypoglycemia are likely to have non-islet cell tumor hypoglycemia. AlthoughIGF-2 protein or mRNA is detectable in tumor tissue, the diagnosis is usually made on clinical grounds, because the plasma levels of IGF-2 are typically not elevated. Levels of IGF binding proteins may be increased.

TREATMENT

The therapy of non-islet cell hypoglycemia should be directed at the underlying cancer: surgical resection or radiation therapy. Patients whose tumors cannot be successfully resected or irradiated can be treated with frequent oral feedings or constant intravenous administration of glucose.

HEMATOLOGIC SYNDROMES

The elevation of granulocyte, platelet, and eosinophil counts in most patients with myeloproliferative disorders is caused by the proliferation of the myeloid elements due to the underlying disease rather than a paraneoplastic syndrome. The paraneoplastic hematologic syndromes in patients with solid tumors are less well characterized than the endocrine syndromes, because the ectopic hormone(s) or cytokines responsible have not been identified in most of these tumors (Table 100-2). The severity of the

paraneoplastic syndromes parallels the course of the cancer.

ERYTHROCYTOSIS

Ectopic production of erythropoietin by cancer cells causes most paraneoplastic erythrocytosis. The ectopically produced erythropoietin stimulates the production of red blood cells in the bone marrow and raises the hematocrit. Other lymphokines and hormones produced by cancer cells may stimulate erythropoietin release but have not been proven to cause erythrocytosis.

Most patients with erythrocytosis have an elevated hematocrit (>52% in men; >48% in women) that is detected on a routine blood count. Approximately 3% of patients with renal cell cancer, 10% of patients with hepatoma, and 15% of patients with cerebellar hemangioblastomas have erythrocytosis. In most cases the erythrocytosis is asymptomatic.

Patients with erythrocytosis due to a renal cell cancer, hepatoma, or central nervous system cancer should have measurement of red cell mass. If the red cell mass is elevated, the serum erythropoietin level should then be measured. Patients with an appropriate cancer, elevated erythropoietin levels, and no other explanation for erythrocytosis (e.g., a hemoglobinopathy that causes increased O_2affinity, Chap. 106) have the paraneoplastic syndrome.

TREATMENT

Successful resection of the cancer usually resolves the erythrocytosis. If the tumor neither can be resected nor treated effectively with radiation therapy or chemotherapy, phlebotomy may control any symptoms related to erythrocytosis.

GRANULOCYTOSIS

Approximately 30% of patients with solid tumors have granulocytosis (granulocyte count>8000/uL). In about half of patients with granulocytosis and cancer, the granulocytosis has an identifiable nonparaneoplastic etiology (infection, tumor necrosis, glucocorticoid administration, etc.). The other patients have proteins in urine and serum that stimulate the growth of bone marrow cells. Tumors and tumor cell lines from patients with lung, ovarian, and bladder cancers have been documented to produce granulocyte colony-stimulating factor (G-CSF), granulocyte-macrophage colony-stimulating factor (GM-CSF), and/orIL-6. However, the etiology of granulocytosis has not been characterized in most patients.

Patients with granulocytosis are nearly all asymptomatic, and the differential white blood cell count does not have a shift to immature forms of neutrophils. Granulocytosis occurs in 40% of patients with lung and gastrointestinal cancers, 20% of patients with breast cancer, 30% of patients with brain tumors and ovarian cancers, and 10% of patients with renal cell carcinoma. Patients with advanced-stage disease are more likely to have granulocytosis than those with early-stage disease.

Paraneoplastic granulocytosis does not require treatment. The granulocytosis resolves

when the underlying cancer is successfully treated.

THROMBOCYTOSIS

Thirty-five percent of patients with thrombocytosis (platelet count>400,000/uL) have an underlying diagnosis of cancer.IL-6, a candidate molecule for the etiology of paraneoplastic thrombocytosis, stimulates the production of platelets in vitro and in vivo. Some patients with cancer and thrombocytosis have elevated levels of IL-6 in plasma. Another candidate molecule is thrombopoietin, a peptide hormone that stimulates megakaryocyte proliferation and platelet production. The etiology of thrombocytosis has not been established in most cases.

Patients with thrombocytosis are nearly all asymptomatic. Thrombocytosis is not clearly linked to thrombosis in patients with cancer. Thrombocytosis is present in 40% of patients with lung and gastrointestinal cancers, 20% of patients with breast, endometrial, and ovarian cancers, and 10% of patients with lymphoma. Patients with thrombocytosis are more likely to have advanced-stage disease and have a poorer prognosis than patients without thrombocytosis. Paraneoplastic thrombocytosis does not require treatment.

EOSINOPHILIA

Eosinophilia is present in ~1% of patients with cancer. Tumors and tumor cell lines from patients with lymphomas or leukemia may produceIL-5, which stimulates eosinophil growth. Activation of IL-5 transcription in lymphomas and leukemias may involve translocation of the long arm of chromosome 5, to which the genes for IL-5 and other cytokines map.

Patients with eosinophilia are typically asymptomatic. Eosinophilia is present in 10% of patients with lymphoma, 3% of patients with lung cancer, and occasional patients with cervical, gastrointestinal, renal, and breast cancer. Patients with markedly elevated eosinophil counts (>5000/uL) can develop shortness of breath and wheezing. A chest radiograph may reveal diffuse pulmonary infiltrates from eosinophil infiltration and activation in the lungs.

TREATMENT

Definitive treatment is directed at the underlying malignancy: tumors should be resected or treated with radiation or chemotherapy. In most patients who develop shortness of breath related to eosinophilia, symptoms resolve with the use of oral or inhaled glucocorticoids.

THROMBOPHLEBITIS

Deep venous thrombosis and pulmonary embolism are the most common thrombotic conditions in patients with cancer. Migratory or recurrent thrombophlebitis may be the initial manifestation of cancer. Approximately 15% of patients who develop deep venous thrombosis or pulmonary embolism have a diagnosis of cancer (Chap. 117). The coexistence of peripheral venous thrombosis with visceral carcinoma, particularly

pancreatic cancer, is called *Trousseau's syndrome*.

Pathogenesis Patients with cancer are predisposed to thromboembolism because they are often at bedrest or immobilized, and tumors may obstruct or slow blood flow. In addition, clotting may be promoted by release of procoagulants or cytokines from tumor cells or associated inflammatory cells, or by platelet adhesion or aggregation. The specific molecules that mediate the increased risk of thromboembolism have not been identified.

Clinical Manifestations Patients with cancer who develop deep venous thrombosis usually develop swelling or pain in the leg, and physical examination reveals tenderness, warmth, and redness. Patients who present with pulmonary embolism develop dyspnea, chest pain, and syncope, and physical examination shows tachycardia, cyanosis, and hypotension. Approximately 5% of patients with no history of cancer who have a diagnosis of deep venous thrombosis or pulmonary embolism will have a diagnosis of cancer within 1 year. The most common cancers associated with thromboembolic episodes include lung, pancreatic, gastrointestinal, breast, ovarian, and genitourinary cancers, lymphomas, and brain tumors. Patients with cancer who undergo surgical procedures requiring general anesthesia have a 20 to 30% risk of deep venous thrombosis.

Diagnosis The diagnosis of deep venous thrombosis in patients with cancer is made by impedance plethysmography or bilateral compression ultrasonography of the leg veins. Patients with a noncompressible venous segment have deep venous thrombosis. If compression ultrasonography is normal and a high clinical suspicion exists for deep venous thrombosis, venography should be done to look for a luminal filling defect. Elevation of D-dimer is not as predictive of deep venous thrombosis in patients with cancer as in patients without cancer.

Patients with symptoms and signs suggesting a pulmonary embolism should be evaluated with a chest radiograph, electrocardiogram, arterial blood gas analysis, and ventilation-perfusion scan. Patients with mismatched segmental perfusion defects have a pulmonary embolus. Patients with equivocal ventilation-perfusion findings should be evaluated as described above for deep venous thrombosis in their legs. If deep venous thrombosis is detected, they should be anticoagulated. If deep venous thrombosis is not detected, they should be considered for a pulmonary angiogram.

Patients without a diagnosis of cancer who present with an initial episode of thrombophlebitis or pulmonary embolus need no additional tests for cancer other than a careful history and physical exam. In light of the many possible primary sites, diagnostic testing in asymptomatic patients is wasteful. However, if the clot is refractory to standard treatment or is in an unusual site, or if the thrombophlebitis is migratory or recurrent, efforts to find an underlying cancer are indicated.

TREATMENT

Patients with cancer and a diagnosis of deep venous thrombosis or pulmonary embolism should be treated initially with intravenous unfractionated heparin or low molecular weight heparin for at least 5 days and coumadin started within 1 or 2 days.

The coumadin dose should be adjusted so the INR is 2 to 3. Patients with proximal deep venous thrombosis and a relative contraindication to heparin anticoagulation (hemorrhagic brain metastases or pericardial effusion) should be considered for placement of a filter in the inferior vena cava (Greenfield filter) to prevent pulmonary embolism. Coumadin should be administered for 3 to 6 months. Patients with cancer who undergo a major surgical procedure should be considered for heparin prophylaxis or pneumatic boots. Breast cancer patients undergoing chemotherapy and patients with implanted catheters should be considered for prophylaxis (1 mg coumadin per day).*Cutaneous paraneoplastic syndromes are discussed in Chap. 57. Neurologic paraneoplastic syndromes are discussed in Chap. 101. More extensive discussion of functional endocrine tumors is given in Chap. 93.*

(Bibliography omitted in Palm version)

Back to Table of Contents

101. PARANEOPLASTIC NEUROLOGIC SYNDROMES - *Muhammad T. Al-Lozi*, *Alan Pestronk*

GENERAL PRINCIPLES

A paraneoplastic neurologic syndrome (PNNS) is a neurologic disorder that is associated with a neoplasm but lies anatomically remote from it. Paraneoplastic disorders are caused by immune or other mechanisms and are not due to direct effects of the tumor itself, metastases, opportunistic infections, complications of drug or radiation therapy, or malnutrition. Clinical features of a PNNS are often distinctive. Onset can be dramatic, arising subacutely over weeks or even days to produce neurologic symptoms that may be profoundly disabling.

PNNSassociated with autoantibodies can be grouped into (1) disorders in which the neoplasm contains a surface antigen or intracellular protein that is the antigenic target, and (2) monoclonal gammopathy syndromes associated with secretion of an antibody by the neoplasm. Each subgroup has typical clinical, pathologic, and immune characteristics (Table 101-1). Some PNNS, including lymphoma-associated motor neuropathy, subacute necrotic myopathy, dermatomyositis, and necrotizing myopathy, have no currently identified antibody or target antigen and are not yet classifiable in this scheme.

The temporal relationship of aPNNS to the associated neoplasm is variable. The PNNS may precede or follow the identification of a neoplasm by weeks, months, or occasionally years. The strength of the association between neoplasms and PNNS varies with different syndromes, different neoplasms, and the clinical context. In some PNNS, such as the sensory neuronopathy associated with anti-Hu antibodies, the association with neoplasm is very strong. By contrast, the Lambert-Eaton myasthenic syndrome (LEMS) is associated with neoplasm in approximately 50% of cases only; the relationship is probably stronger in older individuals who have a history of cigarette smoking. Disorders that are clinically identical to most PNNS also occur in the absence of cancer. Nonetheless, the development of a PNNS in a previously healthy individual should in most circumstances prompt a thorough search for its associated neoplasms.

PNNS-associated neoplasms vary considerably in terms of malignancy. Tumors associated with subacute necrotic myelopathy are often severe and unresponsive to therapy. In other syndromes, the tumor either remains small or can be effectively treated; in such cases, the long-term prognosis is determined by the effectiveness of management of the paraneoplastic syndrome. Thymoma associated with myasthenia gravis is an example of a tumor in this category. In some PNNS, such as those associated with small cell lung cancer (SCLC) and anti-Hu antibodies, it has been suggested that the presence of the autoantibodies may confer a more favorable prognosis by inhibiting tumor growth.

As outlined inTable 101-2, certain syndromes are associated with particular types of tumors, and more than one syndrome may occur with a given neoplasm. For example,SCLCsare associated with a variety ofPNNS, including limbic encephalitis, cerebellar ataxia, opsoclonus-myoclonus, necrotic myelopathy, sensory neuronopathy, autonomic neuropathy, andLEMS.

Prevalence estimates vary with the particular syndrome. Tumors that are most often associated withPNNS are cancers of the lung, stomach, breast, ovary, and colon. Some 30% of patients with thymoma also develop myasthenia gravis as a paraneoplastic syndrome. A poorly characterized neuromyopathy with proximal weakness and distal sensory loss is very common in patients who have lost more than 15% of their body weight. In contrast, most of the well-defined PNNS are rare, with estimated prevalence rates of <1% of the population with cancer.

The diagnosis of aPNNSdepends primarily on (1) the presence of a recognized clinical paraneoplastic syndrome; (2) careful exclusion of other cancer-related disorders; and (3) appropriate confirmatory studies, including measurement of specific antibodies and neurophysiologic studies to define the anatomic distribution of the disease process.

PNNS OF THE CENTRAL NERVOUS SYSTEM

LIMBIC ENCEPHALITIS

Limbic encephalitis is a feature of several paraneoplastic syndromes. It occurs in isolation or overlaps with syndromes that also involve the brainstem, cerebellum, spinal cord, and posterior root ganglia. Patients present with seizures, confusion, psychiatric symptoms (agitation, hallucinations, depression, anxiety, and changes in personality), or severe short-term memory loss. Seizures are commonly complex-partial in type, with or without secondary generalization, and may be intractable and refractory to treatment. Other features can include vertigo, ataxia, nystagmus, numbness/paresthesias, and symmetric or asymmetric weakness. Limbic encephalitis typically progresses over a period of weeks before stabilizing. Exacerbations may occur, and remissions are rare, with or without treatment. The onset of symptoms may precede or follow the discovery of tumor. Limbic encephalitis is most commonly associated withSCLC; less frequently with testicular cancer; and occasionally with thymoma, Hodgkin's disease, non-SCLC, breast, colon, and bladder cancer. Some cases are not associated with cancer.

Magnetic resonance imaging (MRI) suggests that paraneoplastic limbic encephalitis is often a relatively widespread disease of the central nervous system. The T2 signal may be increased not only in the temporal lobes but also in the cortex or brainstem. Approximately 75% of patients show electroencephalographic abnormalities that may include focal slowing and/or paroxysmal sharp waves and spikes. Cerebrospinal fluid (CSF) often shows elevated protein, mild mononuclear pleocytosis, oligoclonal bands, or increased IgG synthesis but may be normal.

Neuropathologic features of limbic encephalitis include neuronal loss in the hippocampus, cingulate gyrus, orbital frontal lobe, brainstem, and posterior root ganglia. Additionally, there may be scattered gliosis, microglial nodules, and/or perivascular lymphocytic cuffing.

Anti-Hu antibodies (Table 101-2) are detected in the serum andCSF in the majority of patients with paraneoplastic limbic encephalitis. Most patients with anti-Hu antibodies haveSCLC; however, breast and prostate cancer and neuroblastoma are also described. Approximately 20% of patients with SCLC without neurologic symptoms have

low titers of anti-Hu antibodies. Anti-Hu antibodies are strongly (>90%) associated with neoplasms; patients with positive antibody titers but a negative initial malignancy workup should have a search for neoplasm repeated every 6 to 12 months. Immunotherapy [plasma exchange, intravenous immunoglobulin (IVIg), cyclophosphamide, or glucocorticoids] and/or resection of primary tumor are only rarely associated with improvement in the limbic encephalitis.

Patients with limbic and/or brainstem encephalitis and testicular cancer may have serum IgG antibodies that bind to Ma2, a 40-kD cytoplasmic and nuclear protein. Ma2 is expressed in both brain tissue and testicular tumors. Occasional patients in this subgroup have improved after treatment of the primary neoplasm.

BRAINSTEM ENCEPHALITIS

Paraneoplastic brainstem encephalitis is usually associated with disease elsewhere in the central or peripheral nervous system. Symptoms of brainstem encephalitis relate to the distribution of the disease process. The predominant symptoms are due to medullary involvement producing nausea, vomiting, nystagmus, vertigo, and ataxia. A rare syndrome of marked dysarthria and dysphagia is associated with pontine involvement. Mesencephalic inflammation and neuronal loss result in nuclear or internuclear eye movement abnormalities; diplopia and oscillopsia may be disabling. Rostral midbrain and nigral involvement may cause rigidity. Other rare disorders are deafness and hypoventilation.

PARANEOPLASTIC CEREBELLAR DEGENERATION (PCD)

Approximately 90% of PCD occurs withSCLC, Hodgkin's lymphoma, or breast or ovarian cancer. Patients usually present with the subacute onset of a pancerebellar disorder consisting of nystagmus, oculomotor ataxia, dysarthric speech, and limb and gait ataxia (Chaps. 22 and364). In many patients, especially those with the anti-Yo antibody syndrome, signs are restricted to cerebellar dysfunction. However, symptoms of more widespread central (lethargy, cognitive abnormalities) and peripheral (weakness, sensory changes, and dry mouth) nervous system involvement may be present. Symptoms usually progress over weeks and eventually stabilize, leaving the patient severely disabled.

MRIusually reveals cerebellar atrophy. TheCSF may be normal or show mildly elevated protein, mononuclear pleocytosis, increased IgG index, and/or oligoclonal bands. The most consistent neuropathologic feature is a diffuse loss of Purkinje cells. Neuronal loss in the granular cell layer and deep cerebellar nuclei may also occur. Perivascular cuffing has been observed in the cerebellum and leptomeninges.

PCDmay be associated with polyclonal IgG anti-Yo, anti-Tr, or anti-glutamate receptor (mGluR1) antibodies in the serum andCSF. Anti-Yo antibodies are commonly associated with breast or ovarian cancer. The Yo autoantigens are proteins that are prominently expressed in Purkinje cell cytoplasm (Golgi) and proximal dendrites but not in the nucleus. Anti-Tr and anti-mGluR1 antibodies occur with Hodgkin's lymphoma. Clinical features in the three antibody groups are similar. PCD syndromes rarely improve after treatment; however, occasional anti-Tr patients improve after treatment of

Hodgkin's lymphoma. Reappearance or exacerbation of the PCD may indicate recurrence of the tumor.

PARANEOPLASTIC OPSOCLONUS-MYOCLONUS (POM)

This disorder is also known as the "dancing eyes-dancing feet" syndrome. Opsoclonic eye movements are involuntary, high-amplitude, arrhythmic, multidirectional, conjugate saccades. They are often nearly continuous and persist with the eyes closed and during sleep. Opsoclonus is associated with blinking and myoclonus and increases with visual pursuit and voluntary ocular refixation. The syndrome may occur in isolation or as a component of otherPNNS, including limbic or brainstem encephalitis. POM occurs in 2% of young children with neuroblastoma and may precede or follow the discovery of the neoplasm; 50% of children with POM harbor neuroblastoma. Patients may also manifest ataxia, irritability, and vomiting. Antibodies directed against neurofilaments have been described. In the pediatric population, POM may improve following treatment with adrenocorticotropic hormone (ACTH), glucocorticoids, orIVIg, but residual central nervous system signs are frequent.

In adults, opsoclonus/myoclonus syndromes may develop in association with neoplasms of the lung (anti-Hu antibodies), breast (anti-Ri antibodies), thymus, lymphoid cells, ovaries, uterus, and bladder. Anti-Ri antibodies bind to neuronal nuclear antigens, including NOVA-1, a protein that regulates RNA splicing or metabolism in a subset of developing neurons. Remissions may occur spontaneously or following treatment of the underlying tumor. Clonazepam and/or valproate may be useful for symptomatic control of opsoclonus and myoclonus.

CARCINOMA-ASSOCIATED RETINOPATHY (CAR)

The chief complaint in CAR is unilateral or bilateral, symmetric or asymmetric, loss of vision. Night blindness may be the presenting symptom. The visual loss occurs either gradually or in a stepwise pattern over weeks to months. Other symptoms may include visual shimmering, sparkling, or distortions. Examination shows poor visual acuity, impaired color vision, and an afferent pupillary defect. Visual field defects most commonly consist of central and/or ring scotomas. CAR occurs mainly withSCLC(90%) but may also occur with melanoma and gynecologic neoplasms. CAR visual loss frequently precedes the discovery of SCLC. CAR associated with melanoma usually follows the discovery of cancer, with an interval of up to 10 years. Histologically, there is severe loss of the inner and outer segments of the rods and cones, with widespread degeneration of the outer nuclear layer. The electroretinogram is usually flat due to loss of the rod and cone cells. Polyclonal IgG antibodies in SCLC/CAR are directed against recoverin, a 23-kD retinal photoreceptor-specific calcium-binding protein. Other autoantigens include retinal enolase, the S-antigen, and tubby-like protein 1 (TULP1) which is a molecule expressed in synaptic terminals of photoreceptor cells. Treatment with glucocorticoids (prednisone) produces mild to moderate improvement in most patients.

PARANEOPLASTIC MYELOPATHY

Paraneoplastic myelopathy is a rare disorder that presents as acute spinal shock

manifest as flaccid paraparesis with a sensory level and sphincter disturbances (Chap. 368). Spinal cord dysfunction is rapidly progressive and ascending. The prognosis is poor. TheCSF is often cellular with a high protein.MRI may show T2 signal changes in the spinal cord, with cord swelling and involvement of both white and gray matter structures. The onset of the syndrome may precede or follow detection of a neoplasm, typically lymphoma, leukemia, or lung cancer.

STIFF-PERSON SYNDROME (SPS)

SPS is characterized by stiffness and painful spasms, especially in axial and proximal limb muscles, due to hyperexcitability of motor neurons (Chap. 22). The stiffness produces lumbar hyperlordosis. Muscle spasms are triggered by stretching, emotion, and sensory stimulation. A small minority of SPS occurs in association with neoplasms such as breast cancer,SCLC, thymoma, Hodgkin's disease, and colon cancer. IgG polyclonal antibodies directed against amphiphysin, a 125-kD synaptic vesicle-associated protein, have been detected in sera of some SPS patients, primarily those with breast cancer. Some patients respond to glucocorticoids or tumor resection. Treatment with diazepam, clonazepam, lioresal, or sodium valproate may produce symptomatic improvement.

PNNS OF THE NEUROMUSCULAR SYSTEM

Paraneoplastic neuromuscular syndromes may selectively involve nerve cell bodies (anterior horn or dorsal root ganglia), peripheral nerves (myelin or motor, sensory, or autonomic axons), the neuromuscular junction, or muscle. Some neuromuscularPNNS, including sensory neuronopathy with anti-Hu antibodies, are almost always associated with cancer. Other syndromes, such as myasthenia gravis, are statistically associated with neoplasms but in up to 90% of patients a neoplasm is never found. Some neuromuscular PNNS have protean clinical manifestations that are not distinctive; correct diagnosis may rely upon serologic testing for specific autoantibodies (Tables 101-2 and101-3).

NEURONOPATHIES

Neuronopathies, indicating damage to the cell body of the neuron, are typically asymmetric and produce proximal as well as distal involvement early in their clinical course. They are generally poorly responsive to treatment.

Subacute sensory neuronopathy (SSN) presents with numbness and pain that evolves in a progressive fashion over 1 to 8 weeks. A history of smoking is found in >95% of patients. Examination shows asymmetric sensory loss that may involve the face and trunk and proximal as well as distal regions of the upper and lower extremities. All sensory modalities are affected. Patients may become disabled by severe sensory ataxia and pseudoathetosis resulting from the deafferentation. Strength is usually normal. Tendon reflexes are diffusely diminished or absent. Nerve conduction studies show diminished or absent sensory responses with normal motor studies. SSN may occur in isolation but is often associated with central nervous system signs, ranging from mild nystagmus to severe encephalopathy.CSFcommonly shows a mild pleocytosis and elevated protein levels. Serum IgG antibodies that bind to the Hu family

of 35- to 40-kDa nuclear proteins are characteristic of SSN and are a useful diagnostic test. Hu proteins are neuron-specific but are also found inSCLCcells. Although anti-Hu antibodies have strong specificity for SSN and otherPNNS, there is no evidence that the antibodies play a pathogenic role. Morphologically there is neuronal loss with perivascular inflammatory infiltrates in the dorsal root ganglia. SSN with anti-Hu antibodies is almost always associated with a neoplasm, especially SCLC, but neoplasm is found at initial evaluation in only 50% of patients. As noted above, the presence of anti-Hu antibodies is associated with a lower degree of malignancy of the SCLC, suggesting that the antibodies may inhibit tumor growth. The differential diagnosis of SSN includes Sjogren's syndrome and drug toxicity from cisplatin or pyridoxine. Treatment consists of therapy for the associated neoplasm. There is almost never improvement in the SSN itself; however, physical therapy may allow the patient, over time, to partially compensate for the sensory loss.

Motor neuronopathy is a rare syndrome that begins subacutely and then reaches a plateau. Patients have asymmetric weakness that may involve the arms more than the legs. The bulbar muscles are spared. Sensation is normal. Motor neuronopathy often manifests after the detection of a neoplasm, typically lymphoma.CSFshows elevated protein levels and oligoclonal bands in 60% of cases.

PERIPHERAL NEUROPATHY

These disorders typically present with distal symptoms and signs in the limbs. They may be axonal, demyelinating, or a combination of the two types. The presence of circulating paraproteins or serologic markers often helps to define paraneoplastic neuropathy syndromes (Table 101-2).

Polyneuropathies associated with circulating paraproteins include: (1) chronic inflammatory demyelinating polyneuropathy (CIDP); (2) demyelinating neuropathy with serum IgM binding to antimyelin-associated glycoprotein (MAG); (3) multifocal motor neuropathy; (4) POEMS syndrome (*p*olyneuropathy, *o*rganomegaly, *e*ndocrinopathy or edema, *M* protein, and *s*kin changes); and (5) primary acquired amyloidosis.**For further discussion of polyneuropathies associated with circulating paraproteins, see Chap. 378.*

An *axonal sensorimotor neuropathy* is frequently associated with neoplasms. This syndrome is especially common in patients with longstanding cancer and substantial weight loss (>15% of baseline weight). The neuropathy is characterized by distal, symmetric sensory loss and paresthesias, which may be painful, and by weakness and muscle wasting, which is especially prominent in the distal legs. Pathologically there is noninflammatory degeneration of axons and mild myelin loss, presumably secondary to the axonopathy. An accompanying myopathy with atrophy of type II muscle fibers may produce proximal muscle weakness. Axonal loss, with low-amplitude sensory and motor amplitudes and normal conduction velocities, is seen on electrophysiologic studies. This neuromyopathy has been described in association with a variety of solid tumors (lung, breast, stomach), lymphoma, and plasma cell dyscrasia. Successful treatment of the neoplasm may result in improvement or stabilization of the neuromyopathy.

Other axonal neuropathies may bePNNS, but their associations with neoplasms are less clearly established. *Peripheral nerve vasculitis*, producing mononeuritis multiplex or

asymmetric sensorimotor polyneuropathy, has been reported with lymphomas or carcinoma of the lung, prostate, kidney, or stomach. *Polyneuropathy*, presenting with either a subacute mononeuritis multiplex or a slowly progressive distal symmetric sensorimotor polyneuropathy, occurs in approximately 20% of patients with cryoglobulinemia. *Guillain-Barre syndrome* (Chap. 378) may be associated with Hodgkin's disease; it is characterized by subacute motor weakness; sensory loss, which is often mild in comparison with the motor deficits; areflexia; and a characteristic elevation of spinal fluid protein concentration without pleocytosis. *Enteric autonomic neuropathy* with anti-Hu antibodies, commonly presenting as intestinal pseudoobstruction, has been described in association withSCLC.

NEUROMUSCULAR JUNCTION

Lambert-Eaton myasthenic syndrome is a disorder of the presynaptic component of neuromuscular transmission. Common symptoms in LEMS are weakness, fatigue, and dryness of the mouth. Some patients complain of paresthesias, myalgia, or impotence. Weakness is symmetric, proximal, and most prominent in the lower limbs. Strength can decrease with rest and improve with exercise. Ocular (diplopia and ptosis) and bulbar (dysphagia and dysarthria) symptoms may occur in some patients. Respiratory muscle weakness is rare. Tendon reflexes are either diminished or absent at rest but may increase after exercise. About 50% of patients have an associated neoplasm, most commonlySCLC and less often a lymphoproliferative disorder. About 3% of patients with SCLC haveLEMS. Almost all patients with both SCLC and LEMS have a smoking history. LEMS may precede the detection of cancer by 2 to 3 years.

The most useful diagnostic test for LEMS is repetitive nerve stimulation, specifically the finding of compound muscle action potential (CMAP) amplitudes that are small at rest but increase by at least 100% after rapid repetitive nerve stimulation (30 to 50 Hz) or maximal muscle contraction sustained for at least 10 s. LEMS is believed to be an autoimmune disorder associated with diminished quantal release of acetylcholine. IgG antibodies directed against P/Q voltage-gated calcium channels (VGCC) in the motor nerve terminal are found in the sera of ~90% of patients with LEMS and in nearly 100% when LEMS is associated with neoplasm. False-positive findings occur with hypergammaglobulinemia, chronic liver disorders, and (in<3% of normal individuals) infection. Ultrastructurally, the number of active zones, which represent the P/Q VGCC in the presynaptic nerve terminal membrane, is decreased in LEMS. This humoral immune response to P/Q VGCC may be stimulated by similar VGCC expressed by the tumor cells. Edrophonium chloride (Tensilon) generally does not improve strength. Treatment of LEMS is directed at tumor resection, enhancing the release of acetylcholine from the presynaptic terminal, and modulating the autoimmune response. Treatment modalities directed at improving neuromuscular transmission include 3,4-diaminopyridine, guanidine, and pyridostigmine. Immunomodulation may include plasmapheresis, glucocorticoids, azathioprine, and cyclosporine.

Myasthenia gravis (MG) is associated with thymoma in 10 to 15% of patients, especially those who present at ³30 years. Fatigable weakness may involve the ocular, facial, bulbar, and/or limb muscles. Anti-acetylcholine receptor antibodies are detected in about 85% of cases. Slow repetitive nerve stimulation (5 Hz) of proximal or facial muscles produces a decrement of³10% in 70% of patients. Single-fiber

electromyography is a sensitive but not specific confirmatory test in difficult cases.**Myasthenia gravis is discussed in Chap. 380.*

MYOPATHY

Mild proximal muscle weakness with type II muscle fiber atrophy is commonly encountered in patients with cancer, especially when weight loss of>15% is present. Muscle wasting is more prominent than muscle weakness in this syndrome. Inflammatory myopathy, especially dermatomyositis in older females, may occur in association with a variety of neoplasms (Chap. 382).

Necrotizing myopathy presents with a subacute onset of weakness that is typically proximal and ranges from mild to severe. Some patients experience myalgia in addition to muscle weakness. The serum creatine kinase levels are very high. Muscle fiber necrosis is the predominant finding in muscle biopsy. Necrotizing myopathy is most commonly seen with adenocarcinoma and non-small cell cancer of the lung but may also be associated with a variety of other neoplasms. The overall prognosis depends on the malignancy of the associated neoplasm. Weakness may improve following tumor resection or glucocorticoid treatment.

Chronic proximal myopathies have been described with an IgM M-protein binding to decorin; scleromyxedema with IgG or IgA M-proteins; a rippling muscle disease has been reported to occur with thymoma. Hormone-secreting (ACTH or parathyroid hormone-like) tumors may also be associated with proximal myopathies.

(Bibliography omitted in Palm version)

Back to Table of Contents

102. ONCOLOGIC EMERGENCIES - *Rasim Gucalp*, *Janice Dutcher*

Emergencies in patients with cancer may be classified into three groups: pressure or obstruction caused by a space-occupying lesion, metabolic or hormonal problems (paraneoplastic syndromes, Chap. 100), and complications arising from the effects of treatment.

STRUCTURAL-OBSTRUCTIVE ONCOLOGIC EMERGENCIES

SUPERIOR VENA CAVA SYNDROME

Superior vena cava syndrome (SVCS) is the clinical manifestation of superior vena cava (SVC) obstruction, with severe reduction in venous return from the head, neck, and upper extremities. Malignant tumors, such as lung cancer, lymphoma, and metastatic tumors, are responsible for more than 90% of all SVCS cases. Lung cancer, particularly of small-cell and squamous-cell histologies, accounts for approximately 85% of all cases of malignant origin. Metastatic cancers to the mediastinum, such as testicular and breast carcinomas, account for a small proportion of cases. Other causes include benign tumors, aortic aneurysm, thyroid enlargement, thrombosis, and fibrosing mediastinitis caused by prior irradiation or histoplasmosis.

Patients withSVCSusually present with neck and facial swelling (especially around the eyes), dyspnea, and cough. Other symptoms include hoarseness, tongue swelling, headaches, nasal congestion, epistaxis, hemoptysis, dysphagia, pain, dizziness, syncope, and lethargy. Bending forward or lying down may aggravate the symptoms. The characteristic physical findings are dilated neck veins, an increased number of collateral veins covering the anterior chest wall, cyanosis, and edema of the face, arms, and chest. More severe cases include proptosis, glossal and laryngeal edema, and obtundation. The clinical picture is milder if the obstruction is located above the azygos vein.

The diagnosis ofSVCS is a clinical one. The most significant chest radiographic finding is widening of the superior mediastinum, most commonly on the right side. Pleural effusion occurs in only 25% of patients, often on the right side. However, a normal chest radiograph is still compatible with the diagnosis if other characteristic findings are present. Computed tomography (CT) provides the most reliable view of the mediastinal anatomy. The diagnosis of SVCS requires diminished or absent opacification of central venous structures with prominent collateral venous circulation. Magnetic resonance imaging (MRI) has no advantages over CT. Invasive procedures, including bronchoscopy, percutaneous needle biopsy, mediastinoscopy, and even thoracotomy, can be performed by a skilled clinician without any major risk of bleeding. For patients with a known cancer, a detailed workup usually is not necessary, and appropriate treatment may be started after obtaining a CT scan of the thorax. For those with no history of malignancy, a detailed evaluation is absolutely necessary to rule out benign causes and determine a specific diagnosis to direct the appropriate therapy.

TREATMENT

The one potentially life-threatening complication of a superior mediastinal mass is

tracheal obstruction. Upper airway obstruction demands emergent therapy. Diuretics with a low salt diet, head elevation, and oxygen may produce temporary symptomatic relief.

Radiation therapy is the primary treatment forSVCScaused by non-small cell lung cancer and other metastatic solid tumors. Chemotherapy is effective when the underlying cancer is small cell carcinoma of the lung or lymphoma. Recurrent SVCS occurs in 10 to 30% of patients after initial therapy; it may be palliated with the use of intravascular self-expanding stents (Fig. 102-1). Surgery may provide immediate relief for patients in whom a benign process is the cause.

Clinical improvement occurs in most patients, although this improvement may be due to the development of adequate collateral circulation. The mortality associated withSVCSdoes not relate to caval obstruction, but rather to the underlying cause.

SVCS and Central Venous Catheters in Adults The use of long-term central venous catheters has become common practice in patients with cancer. Major vessel thrombosis may occur. In these cases, catheter removal should be combined with anticoagulation to prevent embolization. SVCS in this setting, if detected early, can be treated successfully by fibrinolytic therapy without sacrificing the catheter. Warfarin (1 mg/d) reduces the incidence of thrombosis without altering coagulation tests.

PERICARDIAL EFFUSION/TAMPONADE

Malignant pericardial disease is found at autopsy in 5 to 10% of patients with cancer, most frequently with lung cancer, breast cancer, leukemias, and lymphomas. Cardiac tamponade as the initial presentation of extrathoracic malignancy is rare. The origin is not malignancy in about 50% of cancer patients with symptomatic pericardial disease, but can be related to irradiation, drug-induced pericarditis, hypothyroidism, idiopathic pericarditis, infection, or autoimmune diseases. Two types of radiation pericarditis have been described: an acute inflammatory, effusive pericarditis occurring within months of irradiation, which usually resolves spontaneously, and a chronic effusive pericarditis that may appear up to 20 years after radiotherapy and is accompanied by a thickened pericardium.

Most patients with pericardial metastasis are asymptomatic. However, the common symptoms are dyspnea, cough, chest pain, orthopnea, and weakness. Pleural effusion, sinus tachycardia, jugular venous distension, hepatomegaly, peripheral edema, and cyanosis are the most frequent physical findings. Relatively specific diagnostic findings, such as paradoxical pulse, diminished heart sounds, pulsus alternans (pulse waves alternating between those of greater and lesser amplitude with successive beats), and friction rub are less common than with nonmalignant pericardial disease. Chest radiographs and ECG reveal abnormalities in 90% of patients, but half of these abnormalities are nonspecific. Echocardiography is the most helpful diagnostic test. Pericardial fluid may be serous, serosanguineous, or hemorrhagic, and cytologic examination of pericardial fluid is diagnostic in most patients. False negative cytology may occur in patients with lymphoma and mesothelioma.

TREATMENT

Pericardiocentesis with or without the introduction of sclerosing agents, the creation of a pericardial window, complete pericardial stripping, cardiac irradiation, or systemic chemotherapy are effective treatments. Acute pericardial tamponade with life-threatening hemodynamic instability requires immediate drainage of fluid. This can be quickly achieved by pericardiocentesis. Alternatively, subxyphoid pericardiotomy can be performed in 45 min under local anesthesia.

INTESTINAL OBSTRUCTION

Intestinal obstruction and reobstruction are common problems in patients with advanced cancer, particularly colorectal or ovarian carcinoma. However, other cancers, such as lung or breast cancer and melanoma, can metastasize within the abdomen, leading to intestinal obstruction. Typically, obstruction occurs at multiple sites. Intestinal pseudoobstruction is caused by infiltration of the mesentery or bowel muscle by tumor, involvement of the celiac plexus, or paraneoplastic neuropathy in patients with small cell lung cancer. Paraneoplastic neuropathy is associated with IgG antibodies reactive to neurons of the myenteric and submucosal plexuses of the jejunum and stomach. Ovarian cancer can lead either to authentic luminal obstruction or to pseudoobstruction that results when circumferential invasion of a bowel segment arrests the forward progression of peristaltic contractions.

The onset of obstruction is usually insidious. Pain is the most common symptom and is usually colicky in nature. Pain can also be due to abdominal distention, tumor masses, or hepatomegaly. Vomiting can be intermittent or continuous. Patients with complete obstruction usually have constipation. Physical examination may reveal abdominal distention with tympany, ascites, visible peristalsis, high-pitched bowel sounds, and tumor masses. Erect plain abdominal films may reveal multiple air-fluid levels and dilation of the small or large bowel. Acute cecal dilation to more than 12 to 14 cm is considered a surgical emergency because of the high likelihood of rupture. The overall prognosis for the patient with cancer who develops intestinal obstruction is poor; median survival is 3 to 4 months. About one-fourth to one-third of patients are found to have intestinal obstruction due to causes other than cancer. Adhesions from previous operations are a common benign cause. Ileus induced by vincristine is another reversible cause.

TREATMENT

The management of intestinal obstruction in patients with advanced malignancy depends on the extent of the underlying malignancy and the functional status of the major organs. The initial management should include surgical evaluation. Operation is not always successful and may lead to further complications with a substantial mortality rate (10 to 20%). Self-expanding metal stents placed in the gastric outlet, duodenum, proximal jejunum, colon, or rectum may palliate obstructive symptoms at those sites without major surgery. Patients known to have advanced intraabdominal malignancy should receive a prolonged course of conservative management, including nasogastric decompression. Treatment with antiemetics, antispasmodics, and analgesics may allow patients to remain outside the hospital. The somatostatin analogue octreotide may relieve obstructive symptoms through its inhibitory effect on gastrointestinal secretion.

URINARY OBSTRUCTION

Urinary obstruction may occur in patients with prostatic or gynecologic malignancies, particularly cervical carcinoma, or metastatic disease from other primary sites. Radiation therapy to pelvic tumors may cause fibrosis and subsequent ureteral obstruction. Bladder outlet obstruction is usually due to prostate and cervical cancers and may lead to bilateral hydronephrosis and renal failure.

Flank pain is the most common symptom. Persistent urinary tract infection, persistent proteinuria, or hematuria in patients with a cancer should raise suspicion of ureteral obstruction. Total anuria and/or anuria alternating with polyuria may occur. A slow, continuous rise in the serum creatinine level necessitates immediate evaluation in patients with cancer. Renal ultrasound examination is the safest and cheapest way to identify hydronephrosis. The function of an obstructed kidney can be evaluated by a nuclear scan.CT can be helpful in identifying a retroperitoneal mass or retroperitoneal adenopathy.

TREATMENT

Obstruction associated with flank pain, sepsis, or fistula formation is an indication for immediate palliative urinary diversion. There are many newer techniques by which internal ureteral stents can be placed under local anesthesia. Percutaneous nephrostomy offers an alternative approach for drainage. In the case of bladder outlet obstruction due to malignancy, a suprapubic cystostomy can be used for urinary drainage.

MALIGNANT BILIARY OBSTRUCTION

This common clinical problem can be caused by a primary carcinoma arising in the pancreas, ampulla of Vater, bile duct, or liver or by metastatic disease to the periductal lymph nodes or liver parenchyma. The most common metastatic tumors causing biliary obstruction are gastric, colon, breast, and lung cancers. Jaundice, light-colored stools, dark urine, pruritus, and weight loss due to malabsorption are usual symptoms. Pain and secondary infection are uncommon in malignant biliary obstruction. Ultrasound,CT, or percutaneous transhepatic or endoscopic retrograde cholangiography will identify the site and nature of the biliary obstruction.

TREATMENT

Palliative intervention is indicated only in patients with disabling pruritus resistant to medical treatment, severe malabsorption, or infection. Stenting under radiographic control, surgical bypass, or radiation therapy with or without chemotherapy may alleviate the obstruction. The choice of modality should be based on the site of obstruction (proximal versus distal), the type of tumor (sensitive to radiotherapy, chemotherapy, or neither), and the general condition of the patient. In the absence of pruritus, biliary obstruction may be a largely asymptomatic cause of death.

SPINAL CORD COMPRESSION

Spinal cord compression occurs in 5 to 10% of patients with cancer. Epidural tumor is the first manifestation of malignancy in about 10% of patients. The underlying cancer is usually identified during the initial evaluation; lung cancer is most commonly the primary malignancy.

Metastatic tumor involves the vertebral column more often than any other part of the bony skeleton. Lung, breast, and prostate cancer are the most frequent offenders. Multiple myeloma also has a high incidence of spine involvement. The thoracic spine is the most common site (70%), followed by the lumbosacral spine (20%) and the cervical spine (10%). Involvement of multiple sites is most frequent in patients with breast and prostatic carcinoma. Cord injury develops when metastases to the vertebral body or pedicle enlarge and compress the underlying dura. Another cause of cord compression is direct extension of a paravertebral lesion through the intervertebral foramen. These cases usually involve a lymphoma, myeloma, or pediatric neoplasm. Parenchymal spinal cord metastasis due to hematogenous spread is rare.

The most common initial symptom in patients with spinal cord compression is localized back pain and tenderness due to involvement of vertebrae by tumor. Pain is usually present for days or months before other neurologic findings appear. It is exacerbated by movement and by coughing or sneezing. It can be differentiated from the pain of disk disease by the fact that it worsens when the patient is supine. Radicular pain is less common than localized back pain and usually develops later. Radicular pain in the cervical or lumbosacral areas may be unilateral or bilateral. Radicular pain from the thoracic roots is often bilateral and is described by patients as a feeling of tight, band-like constriction around the thorax and abdomen. Typical cervical radicular pain radiates down the arm; in the lumbar region, the radiation is down the legs. Loss of bowel or bladder control may be the presenting symptom, but usually occurs late in the course.

On physical examination, pain induced by straight leg raising, neck flexion, or vertebral percussion may help to determine the level of cord compression. Patients develop numbness and paresthesias in the extremities or trunk. Loss of sensibility to pinprick is as common as loss of sensibility to vibration or position. The upper limit of the zone of sensory loss is often one or two vertebrae below the site of compression. Motor findings include weakness, spasticity, and abnormal muscle stretching. The presence of an extensor plantar reflex reflects significant compression. Deep tendon reflexes may be brisk. Motor and sensory loss usually precede sphincter disturbance. Patients with autonomic dysfunction may present with decreased anal tonus, decreased perineal sensibility, and a distended bladder. The absence of the anal wink reflex or the bulbocavernosus reflex confirms cord (conus or cauda equina) involvement. In doubtful cases, evaluation of post-voiding urinary residual volume can be helpful. A residual volume of more than 150 mL suggests bladder dysfunction. Autonomic dysfunction is an unfavorable prognostic factor. Patients with progressive neurologic symptoms should have frequent neurologic examinations and rapid therapeutic intervention.

Patients with cancer who develop back pain should be evaluated for spinal cord compression as quickly as possible (Fig. 102-2). Treatment is more often successful in patients who are ambulatory and still have sphincter control at the time treatment is

initiated. Patients should have a neurologic examination and plain films of the spine. Those whose physical examination suggests cord compression should receive dexamethasone (24 mg intravenously every 6 h), starting immediately.

Erosion of the pedicles (the "winking owl" sign) is the earliest radiologic finding of vertebral tumor. Other radiographic changes include increased intrapedicular distance, vertebral destruction, lytic or sclerotic lesions, scalloped vertebral bodies, and vertebral body collapse. Vertebral collapse is not a reliable indicator of the presence of tumor; about 20% of cases of vertebral collapse, particularly those in older patients and postmenopausal women, are due not to cancer but to osteoporosis. Also, a normal appearance on plain films of the spine does not exclude the diagnosis of cancer. The role of bone scans in the detection of cord compression is not clear; this method is sensitive but less specific than spinal radiography.

The full-length image of the cord provided byMRI is useful. On T1-weighted images, good contrast is noted between the cord, cerebrospinal fluid, and extradural lesions. Owing to their sensitivity in demonstrating the replacement of bone marrow by tumor, MRI can show which parts of a vertebra are involved by tumor (the body, pedicle, lamina, spinous process). MRI also visualizes intraspinal extradural masses compressing the cord. T2-weighted images are most useful for the demonstration of intramedullary pathology. Gadolinium-enhanced MRI can help to characterize and delineate intramedullary disease. MRI is as good as or better than myelography plus postmyelogramCT in detecting metastatic epidural disease with cord compression. Myelography should be reserved for patients who have poor MR images or who cannot undergo MRI promptly. CT in conjunction with myelography enhances the detection of small areas of spinal destruction.

In patients with spinal cord compression and an unknown primary tumor, a simple workup including chest radiography, mammography, measurement of prostate-specific antigen, and abdominalCTusually reveals the underlying malignancy.

TREATMENT

The treatment of patients with spinal cord compression is aimed at relief of pain and restoration of neurologic function (Fig. 102-2).

Radiation therapy plus glucocorticoids is generally the initial treatment of choice for spinal cord compression. Up to 75% of patients treated when still ambulatory remain ambulatory, but only 10% of patients with paraplegia recover walking capacity. Indications for surgical intervention include unknown etiology, failure of radiation therapy, a radioresistant tumor type (e.g., melanoma or renal cell cancer), pathologic fracture dislocation, and rapidly evolving neurologic symptoms. Until recently, laminectomy was the standard operation for metastatic spinal cord compression, although results were poor. At present, laminectomy should be used only for tissue diagnosis and for the removal of posteriorly localized epidural deposits in the absence of vertebral disease. Because most cases of epidural spinal cord compression are due to anterior or anterolateral extradural disease, resection of the anterior vertebral body along with the tumor, followed by spinal stabilization, has achieved good results and low mortality rate. Chemotherapy may have a role in patients with chemosensitive tumors

who have had prior radiation therapy to the same region and who are not candidates for surgery.

The histology of the tumor is an important determinant of both recovery and survival. Rapid onset and quick progression are poor prognostic features.

INCREASED INTRACRANIAL PRESSURE

About 25% of patients with cancer die with intracranial metastases. The cancers that most often metastasize to the brain are lung and breast cancers and melanoma. Brain metastases often occur in the presence of systemic disease, and they frequently cause major symptoms, disability, and early death.

The signs and symptoms of a metastatic brain tumor are similar to those of other intracranial expanding lesions: headache, nausea, vomiting, behavioral changes, seizures, and focal, progressive neurologic changes. Occasionally the onset is abrupt, resembling a stroke, with the sudden appearance of headache, nausea, vomiting, and neurologic deficits. This picture is usually due to hemorrhage into the metastasis. Melanoma, germ cell tumors, and renal cell cancers have a particularly high incidence of intracranial bleeding. The tumor mass and surrounding edema may cause obstruction of the circulation of cerebrospinal fluid, with resulting hydrocephalus. Patients with increased intracranial pressure may have papilledema with visual disturbances and neck stiffness. As the mass enlarges, brain tissue may be displaced through the fixed cranial openings, producing various herniation syndromes.

CTand MRIare equally effective in the diagnosis of brain metastases. CT with contrast should be used as a screening procedure. The CT scan shows brain metastases as multiple enhancing lesions of various sizes with surrounding areas of low-density edema. If a single lesion or no metastases are visualized by contrast-enhanced CT, MRI of the brain should be performed. Gadolinium-enhanced MRI is more sensitive than CT at revealing small lesions, particularly in the brainstem or cerebellum.

TREATMENT

If signs and symptoms of brain herniation (particularly headache, drowsiness, and papilledema) are present, the patient should be intubated and hyperventilated to maintain P_{CO_2}between 25 and 30 mmHg and should receive infusions of mannitol (1 to 1.5 g/kg) every 6 h. Dexamethasone is the best initial treatment for all symptomatic patients with brain metastases (see above). Patients with multiple lesions should receive whole-brain radiation therapy. Patients with a single brain metastasis and with controlled extracranial disease may be treated with surgical excision followed by whole-brain radiation therapy, especially if they are younger than 60 years. Radioresistant tumors should be resected if possible. Stereotactic radiosurgery is an effective treatment for inaccessible or recurrent lesions. With a gamma knife or linear accelerator, multiple small, well-colimated beams of ionizing radiation destroy lesions seen onMRI. Some patients with increased intracranial pressure associated with hydrocephalus may benefit from shunt placement.

NEOPLASTIC MENINGITIS

Tumor involving the leptomeninges is a complication of both primary tumors of the central nervous system (CNS) and tumors that metastasize to the CNS. The incidence is estimated at 3 to 8% of patients with cancer. Melanoma, breast and lung cancer, lymphoma (including AIDS-associated), and acute leukemia are the most common causes.

Patients typically present with multifocal neurologic signs and symptoms including headache, gait abnormality, mental changes, nausea, vomiting, seizures, back or radicular pain, and limb weakness. Signs include cranial nerve palsies, extremity weakness, paresthesia, and decreased deep tendon reflexes.

Diagnosis is made by demonstrating malignant cells in the cerebrospinal fluid (CSF); however, up to 40% of patients may have false negative CSF cytology. An elevated CSF protein level is nearly always present (except in HLTV-1-associated adult T cell leukemia). Patients with neurologic signs and symptoms consistent with neoplastic meningitis who have a negative CSF cytology but an elevated CSF protein level should have the spinal tap repeated at least three times for repeated cytologic examination before the diagnosis is rejected.MRI may show hydrocephalus or smooth or nodular enhancement of the meninges.

The development of neoplastic meningitis usually occurs in the setting of uncontrolled cancer outside theCNS; thus, prognosis is poor (median survival 10 to 12 weeks). However, treatment of the neoplastic meningitis may successfully alleviate symptoms and control the CNS spread.

TREATMENT

Intrathecal chemotherapy, usually methotrexate, cytarabine, or thiotepa, is delivered by lumbar puncture or by an intraventricular reservoir (Ommaya) three times a week until theCSF is free of malignant cells. Then injections are given twice a week for a month and then weekly for a month. An extended release preparation of cytarabine (Depocyte) has a longer half-life and is more effective than regular formulations. Among solid tumors, breast cancer responds best to therapy. Patients with neoplastic meningitis from either acute leukemia or lymphoma may be cured of theirCNSdisease if the systemic disease can be eliminated.

SEIZURES

Seizures occurring in a patient with cancer can be caused by the tumor itself, by metabolic disturbances, by radiation injury, by cerebral infarctions, by chemotherapy-related encephalopathies, or byCNSinfections. Metastatic disease to the CNS is the most common cause of seizures in patients with cancer. Seizures are a presenting symptom of CNS metastasis in 6 to 29% of cases. Approximately 10% of patients with CNS metastasis eventually develop seizures. The presence of frontal lesions correlates with early seizures, and the presence of hemispheric symptoms increases the risk for late seizures. Both early and late seizures are uncommon in patients with posterior fossa lesions. Seizures are also common in patients with CNS metastases from melanoma. Very rarely, cytotoxic drugs such as etoposide, busulfan,

and chlorambucil cause seizures.

TREATMENT

Patients in whom seizures due toCNSmetastases have been demonstrated should receive anticonvulsive treatment with diphenylhydantoin. Prophylactic anticonvulsant therapy is not recommended unless the patient is at a high risk for late seizures. In those patients, serum diphenylhydantoin levels should be monitored closely and the dosage adjusted accordingly.

INTRACEREBRAL LEUKOCYTOSTASIS

Intracerebral leukocytostasis (Ball's disease) is a potentially fatal complication of acute leukemia (particularly myelogenous leukemia) that can occur when the peripheral blast cell count is greater than 100,000/uL. At such high blast cell counts, blood viscosity is increased and blood flow is slowed, and the primitive leukemic cells are capable of invading through endothelium and causing hemorrhage into the brain. Patients may experience stupor, dizziness, visual disturbances, ataxia, coma, or sudden death. Administration of 600 cGy of whole-brain irradiation can protect against this complication and can be followed by rapid institution of antileukemic therapy. This complication is not a feature of the high white cell counts associated with chronic lymphocytic leukemia or chronic myelogenous leukemia.

HEMOPTYSIS

Hemoptysis may be caused by nonmalignant conditions, but lung cancer accounts for a large proportion of cases. Up to 20% of patients with lung cancer have hemoptysis some time in their course. Endobronchial metastases from carcinoid tumors, breast, colon, kidney cancer, and melanoma may also cause hemoptysis. The volume of bleeding is often difficult to gauge. Massive hemoptysis is defined as more than 600 mL of blood produced in 48 h. When respiratory difficulty occurs, hemoptysis should be treated emergently. Often patients can tell where the bleeding is occurring. They should be placed bleeding side down, given supplemental oxygen, and subjected to emergency bronchoscopy. If the site of the lesion is detected, either the patient undergoes a definitive surgical procedure or the lesion is treated with a neodymium:yttrium-aluminum-garnet (Nd:YAG) laser. The surgical option is preferred. Bronchial artery embolization may control brisk bleeding in 75 to 90% of patients, permitting the definitive surgical procedure to be done more safely. Embolization without definitive surgery is associated with rebleeding in 20 to 50% of patients.

Pulmonary hemorrhage with or without hemoptysis in hematologic malignancies is often associated with fungal infections, particularly *Aspergillus* sp. After granulocytopenia resolves, the lung infiltrates in aspergillosis may cavitate and cause massive hemoptysis. Thrombocytopenia and coagulation defects should be corrected, if possible.

AIRWAY OBSTRUCTION

Generally, *airway obstruction* refers to a blockage at the level of the mainstem bronchi

or above. It may result either from intraluminal tumor growth or from extrinsic compression of the airway. If the obstruction is proximal to the larynx, a tracheostomy may be life-saving. For more distal obstructions, particularly intrinsic lesions incompletely obstucting the airway, bronchoscopy with laser treatment, photodynamic therapy, or stenting can produce immediate relief in most patients. However, radiation therapy (either external-beam irradiation or brachytherapy) given together with glucocorticoids may also open the airway. Symptomatic extrinsic compression may be palliated by stenting.

METABOLIC EMERGENCIES

HYPERCALCEMIA

Hypercalcemia is the most common paraneoplastic syndrome (Chaps. 100 and341), occurring in about 10% of patients with advanced cancer. It is associated most often with cancers of the lung, breast, head and neck, and kidney and with multiple myeloma and some B and T cell lymphomas.

Increased release of calcium from bone is the main factor leading to hypercalcemia. Bone resorption is increased dramatically through stimulation of the proliferation and activity of osteoclasts, and bone formation is not stimulated in parallel. The kidney may play an important role through an increase in the reabsorption of calcium in the distal tubule. Parathormone-related protein (PTHrP) produced by tumors has a central role as a mediator of hypercalcemia in cancer. PTHrP shares 80% homology with the first 13 amino acids of parathormone (PTH), which are in the region responsible for binding to the PTH receptor. PTHrP acts via the PTH hormone receptors on osteoblasts and renal tubular cells to stimulate bone resorption and renal calcium conservation, leading to hypercalcemia. Elevated plasma PTHrP levels are also found in most hypercalcemic patients with bone metastases, whose hypercalcemia has traditionally been explained by local osteolysis due to the production of osteolytic factors by tumors. Transforming growth factors, cytokines (interleukins 1 and 6), and other unknown factors could play a contributory role. True "ectopic" PTH production by malignant tumors is rare. In lymphoma, a vitamin D-related product of the tumor may also increase calcium absorption in the gut.

The clinical features of hypercalcemia in patients with cancer are nonspecific and include fatigue, anorexia, constipation, polydipsia, muscle weakness, nausea, and vomiting. They may easily be attributed to the malignancy itself or to its treatment. Laboratory assessment should include measurement of serum electrolytes, calcium, phosphate, and albumin. Hypoalbuminemia is common in malignancy and affects the total serum concentration of calcium. If the ionized calcium level cannot be obtained, then the corrected serum calcium concentration should be calculated with the following formula:

Most patients with hypercalcemia of malignancy have obvious evidence of malignancy, and their serumPTHlevels are suppressed. Measurements ofPTHrPand serum 1,25-dihydroxyvitamin D are not indicated. Routine serum chemistry evaluations are not

able to distinguish between malignant and nonmalignant causes of hypercalcemia.

TREATMENT

Not all patients with moderate to severe hypercalcemia (corrected calcium³12 mg/dL) should be treated. The decision to treat will depend on the patient's quality of life, the current symptoms, and the prospect for further cancer treatment. Treatment directed at hypercalcemia only extends life in patients for whom effective cancer treatment is available. Nonetheless, therapy may be indicated to reduce symptoms and improve the quality of life. Treatment of symptomatic hypercalcemia begins with intravenous saline to restore the depleted intravascular volume, which may be 4 to 8 L below normal at presentation. Rehydration usually has little effect on calcium levels, producing a median decrease of only 1 mg/dL. Antiresorptive agents are essential to decrease osteoclastic activity and control hypercalcemia. Bisphosphonates, which are potent inhibitors of bone resorption, are easy to administer, virtually free of side effects, and rapidly effective in lowering the serum calcium level. Pamidronate is the most effective of the commercially available bisphosphonates. The recommended dose of pamidronate is 60 mg for moderate hypercalcemia (corrected calcium 12 to 13.5 mg/dL) and 90 mg for severe hypercalcemia (corrected calcium >13.5 mg/dL). The dose is given as a single infusion over 4 or 24 h.

SYNDROME OF INAPPROPRIATE SECRETION OF ANTIDIURETIC HORMONE (SIADH)

SIADH is attributed to production of arginine vasopressin by the tumor cells and is characterized by hyponatremia, urine osmolarity inappropriately higher than plasma osmolarity, and high urinary sodium excretion in the absence of volume depletion. Renal, adrenal, and thyroid insufficiency must be excluded, because these disorders can also present with hyponatremia and impaired urinary dilution. Low serum levels of urea and uric acid are useful in distinguishing SIADH from conditions associated with renal hypoperfusion (Chaps. 100 and329).

A broad spectrum of malignant tumors have been reported to causeSIADH. Ectopic vasopressin secretion may occur in some 38% of small cell carcinomas of the lung; often adrenocorticotropic hormone is also produced. The presence of hyponatremia in patients with small cell lung cancer confers a poor prognosis. SIADH may also be caused by various other conditions, such as CNS and pulmonary disorders and some surgical procedures. A variety of drugs have also been shown to produce SIADH, including antidepressants, angiotensin converting-enzyme inhibitors, and cytotoxic drugs such as vincristine, vinorelbine, ifosfamide, cyclophosphamide, cisplatin, levamisole, and melphalan.

Most patients withSIADHare asymptomatic. The severity of symptoms and signs is related to the degree of hyponatremia and the rapidity with which it develops. Early changes include anorexia, depression, lethargy, irritability, confusion, muscle weakness, and marked personality changes. When the plasma sodium level falls below 110 mEq/L, extensor plantar responses, areflexia, and pseudobulbar palsy may be noted; and further reductions may cause coma, convulsions, and death.

TREATMENT

The optimal therapy forSIADH is to treat the underlying malignancy. If that is not possible, other therapeutic approaches are available, such as water restriction or the administration of demeclocycline (900 to 1200 mg per os bid), urea, or lithium carbonate (300 mg per os tid). Demeclocycline is usually used first. Demeclocycline and lithium inhibit the effects of vasopressin on the distal renal tubule. Patients with seizure or coma from hyponatremia may require normal saline infusion plus furosemide to enhance free water clearance. The rate of sodium correction should be slow [0.5 to 1 (mEq/L)/h] to prevent rapid fluid shifts and central pontine myelinolysis. The serum calcium level should be monitored closely to avoid hypocalcemia.

LACTIC ACIDOSIS

Lactic acidosis is a rare and potentially fatal metabolic complication of cancer. Lactic acidosis associated with sepsis and circulatory failure is a common preterminal event in many malignancies. Lactic acidosis in the absence of hypoxemia may occur in patients with leukemia, lymphoma, or solid tumors. Extensive involvement of the liver by tumor is present in most cases. Alteration of liver function may be responsible for the lactate accumulation. Tachypnea, tachycardia, change of mental status, and hepatomegaly may be seen. The serum level of lactic acid may reach 10 to 20 meq/L (90 to 180 mg/dL). Treatment is aimed at the underlying disease. The danger from lactic acidosis is from the acidosis, not the lactate. Sodium bicarbonate should be added if acidosis is very severe or if hydrogen ion production is very rapid and uncontrolled. The prognosis is poor.

HYPOGLYCEMIA

Persistent hypoglycemia occasionally is associated with tumors other than pancreatic islet cell tumors. Usually these tumors are large, and often they are of mesenchymal origin or are hepatomas or adrenocortical tumors. Mesenchymal tumors are usually located in the retroperitoneum or thorax. In these patients, obtundation, confusion, and behavioral aberrations occur in the postabsorptive period and may precede the diagnosis of the tumor. Hypoglycemia is due to tumor overproduction of insulin-like growth factor, a peptide hormone with structural homology to proinsulin but having only about 1% of its biologic effects. Additionally, the development of hepatic dysfunction from liver metastases and increased glucose consumption by the tumor can contribute to hypoglycemia. If the tumor cannot be resected, treatment of the hypoglycemia has generally been relief of symptoms, with the administration of glucose, glucocorticoids, or glucagon.

Hypoglycemia can be artifactual; hyperleukocytosis from leukemia, myeloproliferative diseases, leukemoid reactions, or colony stimulating factor treatment can increase glucose consumption in the test tube after blood is drawn, leading to pseudohypoglycemia.

ADRENAL INSUFFICIENCY

In patients with cancer, adrenal insufficiency may go unrecognized because the

symptoms, such as nausea, vomiting, anorexia, and orthostatic hypotension, are nonspecific and may be mistakenly attributed to progressive cancer or to cancer therapy. Primary adrenal insufficiency may develop owing to replacement of both glands by metastases (lung, breast, colon, or kidney cancer, lymphoma), to removal of both glands, or to hemorrhagic necrosis in association with sepsis or anticoagulation. Impaired adrenal steroid synthesis occurs in patients being treated for cancer with mitotane, ketoconazole, aminoglutethimide, or the investigational agent suramin or in those undergoing rapid reduction in glucocorticoid therapy. Rarely, metastatic replacement causes primary adrenal insufficiency as the first manifestation of an occult malignancy. Metastasis to the pituitary or hypothalamus is found at autopsy in up to 5% of patients with cancer, but associated secondary adrenal insufficiency is rare. Patients abruptly discontinuing megestrol acetate therapy (for cancer cachexia) may develop Addisonian crisis from central suppression of the pituitary-adrenal axis with decreased serum levels of cortisol and adrenocorticotropic hormone.

Acute adrenal insufficiency is potentially lethal. Treatment of suspected adrenal crisis is initiated after the sampling of serum cortisol and ACTH levels (Chap. 331).

TREATMENT-RELATED EMERGENCIES

TUMOR LYSIS SYNDROME

Tumor lysis syndrome is a well-recognized clinical entity that is characterized by various combinations of hyperuricemia, hyperkalemia, hyperphosphatemia, lactic acidosis, and hypocalcemia and is caused by the destruction of a large number of rapidly proliferating neoplastic cells. Frequently, acute renal failure develops as a result of the syndrome.

Tumor lysis syndrome is most frequently associated with the treatment of Burkitt's lymphoma, acute lymphoblastic leukemia, and other high-grade lymphomas, but it also may be seen with chronic leukemias and, rarely, with solid tumors. This syndrome has been seen in patients with chronic lymphocytic leukemia after treatment with fludarabine and cladribine. Tumor lysis syndrome usually occurs during or shortly (1 to 5 days) after chemotherapy. Rarely, spontaneous necrosis of malignancies causes tumor lysis syndrome.

Hyperuricemia may be present at the time of chemotherapy. Effective treatment accelerates the destruction of malignant cells and leads to increased serum uric acid levels from the turnover of nucleic acids. Owing to the acidic local environment, uric acid can precipitate in the tubules, medulla, and collecting ducts of the kidney, leading to renal failure. Lactic acidosis and dehydration may contribute to the precipitation of uric acid in the renal tubules. The finding of uric acid crystals in the urine is strong evidence for uric acid nephropathy. The ratio of urinary uric acid to urinary creatinine is >1 in patients with acute hyperuricemic nephropathy and<1 in patients with renal failure due to other causes.

Hyperphosphatemia, which can be caused by the release of intracellular phosphate pools by tumor cell lysis, produces a reciprocal depression in serum calcium, which causes severe neuromuscular irritability and tetany. Deposition of calcium phosphate in the kidney and hyperphosphatemia may cause renal failure. Potassium is the principal

intracellular cation, and massive destruction of malignant cells may lead to hyperkalemia. Hyperkalemia in patients with renal failure may rapidly become life-threatening. Hyperkalemia can cause ventricular arrhythmias and sudden death.

The likelihood that the tumor lysis syndrome will occur in patients with Burkitt's lymphoma is related to the tumor burden and renal function. Hyperuricemia and high serum levels of lactate dehydrogenase LDH (>1500 U/L), both of which correlate with total tumor burden, also correlate with the risk of tumor lysis syndrome. In patients at risk for tumor lysis, pretreatment evaluations should include a complete blood count, serum chemistry evaluation, and urine analysis. High leukocyte and platelet counts may artificially elevate potassium levels ("pseudohyperkalemia") due to lysis of these cells after the blood is drawn. In these cases, plasma potassium instead of serum potassium should be followed. In pseudohyperkalemia, no electrocardiographic abnormalities are present. In patients with abnormal baseline renal function, the kidneys and retroperitoneal area should be evaluated by sonography and/orCT. Urine output should be watched closely.

Recognition of risk and prevention are the most important steps in the management of this syndrome (Fig. 102-3). Despite aggressive prophylaxis, tumor lysis syndrome and/or oliguric or anuric renal failure may occur. Dialysis is often necessary and should be considered early in the course. Hemodialysis is preferred. The prognosis is excellent, and renal function recovers after the uric acid level is lowered to <10 to 20 mg/dL.

HUMAN ANTIBODY INFUSION REACTIONS

The initial infusion of human or humanized antibodies (e.g., rituximab) is associated with fever, chills, nausea, asthenia, and headache in up to half of treated patients. Bronchospasm and hypotension occur in 1% of patients. The pathogenesis is thought to be activation of immune effector processes (cells and complement). In the presence of high levels of circulating tumor cells, thrombocytopenia, a rapid fall in circulating tumor cells, and mild tumor lysis syndrome may also occur. Diphenhydramine and acetaminophen can often prevent or suppress the symptoms. If they occur, the infusion should be stopped and restarted at half the initial infusion rate after the symptoms have abated.

HEMOLYTIC-UREMIC SYNDROME

Hemolytic-uremic syndrome (HUS) and, less commonly, thrombotic thrombocytopenic purpura (TTP) occurring after treatment with antineoplastic drugs have been described. Mitomycin is by far the most common agent causing this peculiar syndrome. Other chemotherapeutic agents, including cisplatin, bleomycin, and gemcitabine, have also been reported to be associated with this syndrome. It occurs most often in patients with gastric, colorectal, and breast carcinoma. In one series, 35% of patients were without evident cancer at the time this syndrome appeared. Secondary HUS/TTP has also been reported as a rare but sometimes fatal complication of bone marrow transplantation.

HUSusually has its onset 4 to 8 weeks after the last dose of chemotherapy, but it is not rare to detect it several months later. HUS is characterized by microangiopathic hemolytic anemia, thrombocytopenia, and renal failure. Dyspnea, weakness, fatigue,

oliguria, and purpura are also common initial symptoms and findings. Systemic hypertension and pulmonary edema frequently occur. Severe hypertension, pulmonary edema, and rapid worsening of hemolysis and renal function may occur after a blood transfusion. Cardiac findings include atrial arrhythmias, pericardial friction rub, and pericardial effusion. Raynaud's phenomenon is part of the syndrome in patients treated with bleomycin.

Laboratory findings include severe to moderate anemia associated with red blood cell fragmentation and numerous schistocytes on peripheral smear. Reticulocytosis, decreased plasma haptoglobin, and an elevated lactic dehydrogenase (LDH) level document hemolysis. The serum bilirubin level is usually normal or slightly elevated. The Coombs test is negative. The white cell count is usually normal, and thrombocytopenia (<100,000/uL) is almost always present. Most patients have a normal coagulation profile, although some have mild elevations in thrombin time and in level of fibrin degradation products. The serum creatinine level is elevated at presentation and shows a pattern of subacute worsening within weeks of the initial azotemia. The urinalysis reveals hematuria, proteinuria, and granular or hyaline casts; and circulating immune complexes may be present.

The basic pathologic lesion appears to be deposition of fibrin in the walls of capillaries and arterioles, and these deposits are similar to those seen inHUS due to other causes. These microvascular abnormalities involve mainly the kidneys and rarely occur in other organs. The pathogenesis of chemotherapy-related HUS is unknown. Immune complexes have been proposed but not confirmed to be etiologic.

The case fatality rate is high; most patients die within a few months. Plasmapheresis and plasma exchange may normalize the hematologic abnormalities, but renal failure is not reversed in most patients. Immunoperfusion over a staphylococcal protein A column is the most successful treatment. About half of the patients treated with immunoperfusion respond with resolution of thrombocytopenia, improvement in anemia, and stabilization of renal failure. Treatment is well tolerated. It is not clear how the treatment works.

NEUTROPENIA AND INFECTION

These remain the most common serious complications of cancer therapy.**They are covered in detail in Chap. 85.*

PULMONARY INFILTRATES

Patients with cancer may present with dyspnea associated with diffuse interstitial infiltrates on chest radiographs. Such infiltrates may be due to progression of the underlying malignancy, treatment-related toxicities, infection, and/or unrelated diseases. The cause may be multifactorial; however, most commonly they occur as a consequence of treatment. Infiltration of the lung by malignancy has been described in patients with leukemia, lymphoma, and breast and other solid cancers. Pulmonary lymphatics may be involved diffusely by neoplasm (pulmonary lymphangitic carcinomatosis), resulting in a diffuse increase in interstitial markings on chest radiographs. The patient is often mildly dyspneic at the onset, but pulmonary failure

develops over a period of weeks. In some patients, dyspnea precedes changes on the chest radiographs and is accompanied by a nonproductive cough. This syndrome is characteristic of solid tumors. In patients with leukemia, diffuse microscopic neoplastic peribronchial and peribronchiolar infiltration is frequent but may be asymptomatic. However, some patients present with diffuse interstitial infiltrates, an alveolar capillary block syndrome, and respiratory distress. In these situations, glucocorticoids can provide symptomatic relief, but specific chemotherapy should always be started promptly.

In addition, accumulation of leukemic blasts in the pulmonary capillary system may cause pulmonary distress and failure in patients with acute myelogenous leukemia. This complication is strongly related to high peripheral blast counts (>100,000/uL) and a short tumor cell doubling time. Some patients with pulmonary leukostasis may have nodular and/or floccular, diffuse infiltrates on chest radiographs. In addition to dyspnea, patients may develop dizziness, confusion, tinnitus, ataxia, visual blurring, and retinal abnormalities due to leukostasis in cerebral vessels. Leukapheresis and/or chemotherapy should be started without delay. Pulmonary irradiation may reduce symptoms.

Several cytotoxic agents, such as bleomycin, methotrexate, busulfan, and the nitrosoureas, may cause pulmonary damage. The most frequent presentations are interstitial pneumonitis, alveolitis, and pulmonary fibrosis. Some cytotoxic agents, including methotrexate and procarbazine, may cause an acute hypersensitivity reaction. Cytosine arabinoside has been associated with noncardiogenic pulmonary edema. Administration of multiple cytotoxic drugs, as well as radiation therapy and preexisting lung disease, may potentiate the pulmonary toxicity. Supplemental oxygen may potentiate the effects of drugs and radiation injury. Patients should always be managed with the lowest FI_{O2} that is sufficient to maintain hemoglobin saturation.

The onset of symptoms may be insidious, with symptoms including dyspnea, nonproductive cough, and tachycardia. Patients may have bibasilar crepitant rales, end-inspiratory crackles, fever, and cyanosis. The chest radiograph generally shows an interstitial and sometimes an intraalveolar pattern that is strongest at the lung bases and may be symmetric. A small effusion may occur. Hypoxemia with decreased carbon monoxide diffusing capacity is always present. Glucocorticoids may be helpful in patients in whom pulmonary toxicity is related to radiation therapy or to chemotherapy. Treatment is otherwise supportive.

Radiation pneumonitis and/or fibrosis is a relatively frequent side effect of thoracic radiation therapy when the dosage exceeds 40 Gy; it may be acute or chronic. It has its onset usually from 2 to 6 months after completion of radiation therapy. The clinical syndrome, which varies in severity, consists of dyspnea, cough with scanty sputum, low-grade fever, and an initial hazy infiltrate on chest radiographs. The infiltrate and tissue damage generally are confined to the radiation field. The patients subsequently may develop a patchy alveolar infiltrate and air bronchograms, which may progress to acute respiratory failure that is sometimes fatal. A lung biopsy may be necessary to make the diagnosis. Asymptomatic infiltrates found incidentally after radiation therapy need not be treated. However, prednisone should be administered to patients with fever or other symptoms. The dosage should be tapered slowly after the resolution of

radiation pneumonitis, as abrupt withdrawal of glucocorticoids may cause an exacerbation of pneumonia. Delayed radiation fibrosis may occur years after radiation therapy and is signaled by dyspnea on exertion. Often it is mild, but it can progress to chronic respiratory failure. Therapy is supportive.

Classical radiation pneumonitis that leads to pulmonary fibrosis is due to radiation-induced production of local cytokines such as platelet-derived growth factor b, tumor necrosis factor, and transforming growth factorb in the radiation field. An immunologically mediated sporadic radiation pneumonitis occurs in about 10% of patients; bilateral alveolitis mediated by T cells results in infiltrates outside the radiation field. This form of radiation pneumonitis usually resolves without sequelae.

Pneumonia is a common problem in patients undergoing treatment for cancer. Bacterial pneumonia typically causes a localized infiltrate on chest radiographs. Therapy is tailored to the causative organism. When diffuse interstitial infiltrates appear in a febrile patient, the differential diagnosis is extensive and includes pneumonia due to infection with *Pneumocystis carinii*, cytomegalovirus, or intracellular pathogens such as mycoplasma and *Legionella*; effects of drugs or radiation; tumor progression; nonspecific pneumonitis; and fungal disease. Patients with cancer who are neutropenic and have fever and local infiltrates on chest radiograph should be treated with a third generation cephalosporin perhaps together with an aminoglycoside or imipenem. A new or persistent focal infiltrate not responding to broad spectrum antibiotics argues for initiation of empiric antifungal therapy. When diffuse bilateral infiltrates develop in patients with febrile neutropenia, broad spectrum antibiotics plus trimethoprim-sulfamethoxazole with or without erythromycin should be initiated. The empiric administration of trimethoprim-sulfamethoxazole plus erythromycin to patients without neutropenia and these antibiotics plus ceftazidime to patients with neutropenia covers nearly every treatable diagnosis (except tumor progression) and gives as good overall survival as a strategy based on early invasive intervention with bronchoalveolar lavage or open lung biopsy. If the patient does not improve in 4 days, open lung biopsy is the procedure of choice. Bronchoscopy with bronchoalveolar lavage may be used in patients who are poor candidates for surgery.

In patients with pulmonary infiltrates who are afebrile, heart failure and multiple pulmonary emboli form part of the differential diagnosis.

TYPHLITIS

Neutropenic enterocolitis (typhlitis) is a necrosis of the cecum and adjacent colon that may complicate the treatment of acute leukemia. The patient develops right lower quadrant abdominal pain, often with rebound tenderness and a tense, distended abdomen, in a setting of fever and neutropenia. Watery diarrhea (often containing sloughed mucosa) and bacteremia are common, and bleeding may occur. Plain abdominal films are generally of little value in the diagnosis;CT scan may show marked bowel wall thickening, particularly in the cecum, with bowel wall edema. Rapid institution of broad-spectrum antibiotic coverage and nasogastric suction may reverse the disease. Surgical intervention should be considered if there is no improvement by 24 h after the start of antibiotic treatment. If the localized abdominal findings become diffuse, the prognosis is poor.

HEMORRHAGIC CYSTITIS

Hemorrhagic cystitis can develop in patients receiving cyclophosphamide or ifosfamide. Both drugs are metabolized to acrolein, which is a strong chemical irritant that is excreted in the urine. Prolonged contact or high concentrations may lead to bladder irritation and hemorrhage. Symptoms include gross hematuria, frequency, dysuria, burning, urgency, incontinence, and nocturia. The best management is prevention. Maintaining a high rate of urine flow minimizes exposure. In addition, 2-mercaptoethanesulfonate (mesna) detoxifies the metabolites and can be coadministered with the instigating drugs. Mesna usually is given three times on the day of ifosfamide administration in doses that are each 20% of the total ifosfamide dose. If hemorrhagic cystitis develops, the maintenance of a high urine flow may be sufficient supportive care. If conservative management is not effective, irrigation of the bladder with an 0.37 to 0.74% formalin solution for 10 min stops the bleeding in most cases. *N*-acetylcysteine may also be an effective irrigant. Prostaglandins (carboprost tromethamine) can inhibit the progress. In extreme cases, ligation of the hypogastric arteries, urinary diversion, or cystectomy may be necessary.

In summary, the diagnosis of cancer and its treatment carry risk of a multitude of medical problems. Knowledge of both the disease process and the potential hazards of the treatment is required to anticipate and treat these emergent complications.

(Bibliography omitted in Palm version)

Back to Table of Contents

103. LATE CONSEQUENCES OF CANCER AND ITS TREATMENT - *Michael C. Perry*, *Dan L. Longo*

The 5-year survival rate of all patients diagnosed with cancer is now 59%. This year alone, nearly 700,000 survivors will be added to the 7 million already considered cured. Virtually all of these survivors will bear some mark of their diagnosis and its therapy, and many will experience long-term complications, including medical problems, psychosocial disturbances, sexual dysfunction, and inability to find employment or insurance.

Problems may be related to the cancer itself (for example, patients with primary cancers of the head and neck are at increased risk for subsequent lung cancer) or to the normal aging process (surviving one cancer does not necessarily alter the risk of other common tumors that increase in frequency with age). However, many of the problems affecting cured patients are related to the treatments. Large numbers of individuals carefully followed for periods up to 30 years have taught us the spectrum of problems that can be encountered. Because of heterogeneity in treatment details and in completeness of follow-up, some treatment-related problems went undetected for many years. However, studies of long-term survivors of childhood cancers, acute leukemia, Hodgkin's disease, lymphomas, testicular cancer, and localized solid tumors have identified the features of cancer treatment that are associated with later morbidity and mortality. We have been somewhat slow to act in changing those aspects of primary treatment that contribute to these late problems. This reticence is due to the uncertainty associated with changing a treatment that is known to work before having a replacement that works as well.

The first task is always to eradicate the diagnosed malignancy. Late problems occurring in cured patients reflect the success of treatment. Such problems never develop in those who do not survive the cancer. Morbidity and mortality from iatrogenic disease should be avoided, if possible. However, the risk of late complications should not lead to the failure to apply potentially curative treatment. The challenge is to preserve or augment the cure rate while decreasing the risk of serious treatment-related illness.

The mechanisms of damage vary. Surgical procedures can create abnormal physiology (such as blind loops leading to malabsorption) or interfere with normal organ function (splenectomy leading to impaired immune response). Radiation therapy can damage organ function directly (salivary gland toxicity leading to dry mouth and dental caries), act as a carcinogen (second solid tumors in radiation ports), or promote accelerated aging-associated changes (atherosclerosis). Cancer chemotherapy can produce damage to the bone marrow and immune system and induce a spectrum of organ dysfunctions. Therapy may produce subclinical damage that may only become recognized in the presence of a second inciting factor (such as the increased incidence of melanoma in patients with dysplastic nevus syndrome treated for Hodgkin's disease with radiation therapy). Finally, although the mechanisms are not elucidated, cancer and its treatment are associated with psychosocial problems that can impair the survivor's ability to adapt to life after cancer.

Late effects by treatment modality are shown in Table 103-1.**Toxicities associated with drugs are discussed in Chap. 84; radiation toxicity is discussed in Chap 394.*

CONSEQUENCES BY ORGAN SYSTEM

Cardiovascular Dysfunction Most anthracyclines damage the heart muscle. A dose-dependent dropout of myocardial cells is seen on endomyocardial biopsy, and eventually ventricular failure ensues. About 5% of patients who receive>550 mg/m_2 of doxorubicin will develop congestive heart failure (CHF). Coexisting cardiac disease, hypertension, advanced age, and concomitant therapy with thoracic radiation therapy or mitomycin may hasten the onset of CHF. Anthracycline-induced CHF is not readily reversible; mortality is as high as 50%, thus, prevention is the best approach. Mitoxantrone is a related drug that has less cardiac toxicity. Administration of doxorubicin by continuous infusion or encapsulated in liposomes appears to decrease the risk of heart damage. Dexrazoxane, an intracellular iron chelator, may protect the heart against anthracycline toxicity by preventing iron-dependent free-radical generation.

Mediastinal radiation therapy that includes the heart can induce acute pericarditis, chronic constrictive pericarditis, myocardial fibrosis, or accelerated premature coronary atherosclerosis. The incidence of acute pericarditis is 5 to 13%; patients may be asymptomatic or have dyspnea on exertion, fever, chest pain. Onset is insidious with a peak about 9 months after treatment. Pericardial effusion may be present. Chronic constrictive pericarditis can develop 5 to 10 years after treatment and usually presents with dyspnea on exertion. Myocardial fibrosis may present as unexplainedCHF with diagnostic evaluation showing restrictive cardiomyopathy. Patients may have aortic insufficiency from valvular thickening or mitral regurgitation from papillary muscle dysfunction. Patients who receive mantle field radiation therapy have a three-fold increased risk of *fatal* myocardial infarction. Similarly, radiation of the carotids is associated with premature atherosclerosis of the carotids and can produce central nervous system (CNS) embolic disease.

At very high doses, cyclosphosphamide can produce a hemorrhagic myocarditis. Patients receiving bleomycin may develop Raynaud's phenomenon. The symptoms range from mild to debilitating; up to 40% of patients receiving bleomycin for testicular cancer report this problem.

Pulmonary Dysfunction Pulmonary fibrosis from bleomycin is dose-related, with potential exacerbation by age, preexisting lung disease, thoracic radiation, high concentrations of inhaled oxygen, and the concomitant use of other chemotherapeutic agents. Several other chemotherapy agents and radiation therapy can cause pulmonary fibrosis, and at least five can cause pulmonary venoocclusive disease, especially following high-dose therapy such as that involved in stem cell/bone marrow transplantation.

Liver Dysfunction Clinically significant long-term damage to the liver from standard dose chemotherapy is relatively infrequent, and mostly confined to patients who have received chronic methotrexate for maintenance therapy of acute lymphoblastic leukemia. Radiation doses to the liver exceeding 1500 cGy can produce liver dysfunction. Although rarely seen with standard dose chemotherapy, hepatic venoocclusive disease is more common with high-dose therapy, such as that given to prepare patients for autologous or allogeneic stem cell transplantation. Endothelial damage is probably the inciting event.

Renal/Bladder Dysfunction Reduced renal function may be produced by cisplatin and is usually asymptomatic, but may also render the patient that much more susceptible to other renal insults. Cyclophosphamide cystitis may eventually lead to the development of bladder cancer. Ifosfamide produces cystitis and a proximal tubular defect, a Fanconi-like syndrome that is usually, but not always, reversible.

Endocrine Dysfunction Long-term survivors of childhood cancer who received cranial irradiation are shorter, more likely to be obese, and have reductions in strength, exercise tolerance, and bone mineral density. The obesity may be related to alterations in leptin biology. Growth hormone deficiency is the most common hormone deficiency.

Thyroid disease is common in patients who have received radiation therapy to the neck, such as patients with Hodgkin's disease, with an incidence of up to 62% at 26 years post-therapy. Hypothyroidism is the most common abnormality, followed by Graves' disease, thyroiditis, and cancer. Such patients should have frequent thyroid-stimulating hormone (TSH) levels to detect hypothyroidism early and suppress the TSH drive, which may contribute to thyroid cancer.

Nervous System Dysfunction Although many patients experience peripheral neuropathy during chemotherapy, only a few have chronic problems, perhaps because they have other co-existing diseases such as diabetes mellitus. High doses of cisplatin can produce severe sensorimotor neuropathy. Vincristine may produce permanent numbness and tingling in the fingers and toes.

Neurocognitive sequelae from intrathecal chemotherapy, with or without radiation therapy, are recognized complications of the successful therapy of childhood acute lymphoblastic leukemia. Cognitive decline has been attributed to radiating the brain in the treatment of a variety of tumor types. In addition, cognitive decline can follow the use of adjuvant chemotherapy in women being treated for breast cancer. Because the agents are given at modest doses and are not thought to cross the blood-brain barrier, the mechanism of the cognitive decline is not defined.

Many patients suffer intrusive thoughts about cancer recurrence for many years after successful treatment. Adjustment to normal expectations can be difficult. Cancer survivors may often have more problems holding a job, staying in a stable relationship, and coping with the usual stresses of daily life.

A dose-related hearing loss can occur with the use of cisplatin, usually with doses in excess of 400 mg/m_2. This is irreversible and patients should be screened with audiometric exams periodically during such therapy.

Eyes Cataracts may be caused by chronic glucocorticoid use, radiation therapy to the head, and, rarely, by tamoxifen.

Sexual and Reproductive Dysfunction Reversible azoospermia can be caused by many chemotherapy agents. The gonads may also be permanently damaged by radiation therapy or by chemotherapeutic agents, particularly the alkylating agents. The extent of the damage depends upon the patient's age and the total dose administered.

As a woman nears menopause, smaller amounts of chemotherapy will produce ovarian failure. In men, chemotherapy may produce infertility, but hormone production is not usually affected. Women, however, commonly lose both fertility and hormone production. The premature induction of menopause in a young woman can have serious medical and psychological consequences. Hormone replacement therapy is controversial, but most evidence supports its use. Paroxetine may be useful in controlling hot flashes.

Musculoskeletal Dysfunction Late consequences of radiation therapy on the musculoskeletal system occur mostly in children and are related to the radiation dose, volume of tissue irradiated, and the age of the child at the time of therapy. Damage to the microvasculature of the epiphyseal growth zone may result in leg length discrepancy, scoliosis, and short stature.

Oral Complications Radiation therapy can damage the salivary glands, producing dry mouth. Without saliva, dental caries develop and many patients have poor dentition. In rare patients, taste can be adversely affected and appetite can be suppressed.

SECOND MALIGNANCIES

Second malignancies are a major cause of death for those cured of cancer. Second malignancies can be grouped into three categories: those associated with the primary cancer, those caused by radiation therapy, and those caused by chemotherapy.

Primary cancers increase the risk of secondary cancers in a number of settings. Patients with head and neck cancers are at increased risk of developing a lung cancer, and vice versa, probably because of shared risk factors, especially tobacco abuse. Patients with breast cancer are at increased risk of a second breast cancer in the contralateral breast. Patients with Hodgkin's disease are at increased risk of non-Hodgkin's lymphoma. Patients with genetic syndromes, such as MEN 1 or Lynch syndrome, are at increased risk of second cancers of specific types. In none of these examples does it appear that treatment of primary cancer is the cause of the secondary cancer, but a role for treatment is difficult to exclude. These predispositions should result in heightened surveillance in persons at risk. Patients with head and neck cancer may have a reduced risk of developing lung cancer with retinoic acid treatment. Other cancer preventions have not been proved effective.

Patients treated with radiation therapy have an increasing and apparently life-long risk of developing second solid tumors, usually in or adjacent to the radiation field. The risk is modest in the first decade after treatment but reaches 1% per year in the second decade, such that populations followed for 25 years or more have a³25% chance of developing a second treatment-related tumor. Some organs differ in their susceptibility to radiation carcinogenesis with age; women receiving chest radiation therapy after age 30 have a small increased risk of breast cancer, but those under 30 have a 128-fold increased risk. The chances of curing the second malignancies hinge on early diagnosis. Patients who were treated with radiation therapy should be carefully examined on an annual basis and evaluated for any abnormalities in organs and tissues that were in the radiation field. Symptoms in a patient cured of cancer should not be dismissed as they may be an early sign of second cancers.

Chemotherapy produces two clinical syndromes that can be fatal: myelodysplasia and acute myeloid leukemia. Two types of acute leukemia have been described. The first occurs in patients treated with alkylating agents, especially over a protracted period. The malignant cells frequently carry genetic deletions in chromosomes 5 or 7. The lifetime risk is about 2%; the risk is increased by the addition of radiation therapy and is about 3 times higher in people treated over age 40. It peaks in incidence 4 to 6 years after treatment; the risk returns to baseline if no disease has developed within 10 years of treatment. The second type of acute leukemia occurs after exposure to topoisomerase II inhibitors such as doxorubicin or etoposide. It is morphologically indistinguishable from the first but contains a characteristic chromosome translocation involving 10q23. The incidence is <1%, and it usually occurs 1 1/2 to 3 years after treatment. Both forms of acute leukemia are highly refractory to treatment, and no preventive strategy has been developed.

Hormonal manipulations can also cause second tumors. Tamoxifen induces endometrial cancer in about 1 to 2% of women taking it 5 years or longer. Usually these tumors are found at early stage; mortality from endometrial cancer is very low compared to the benefit from tamoxifen use as adjuvant therapy in women with breast cancer.

CONSEQUENCES BY CANCER TYPE

Pediatric Cancers Quality of life is often excellent, although the majority have at least one late effect. About one-third of long-term survivors have moderate to severe problems. Cognitive function may be impaired. Late effects are worse for those with poor socioeconomic status. Functional impairments in the cardiovascular system due to radiation therapy and anthracyclines, and in the lungs due to radiation therapy, are rare. Scoliosis and/or delayed growth due to radiation of the skeleton is more common. Many have psychosocial and sexual problems. Second malignant neoplasms are a significant cause of death.

Hodgkin's Disease The patient cured of Hodgkin's disease remains subject to long-term medical problems such as thyroid dysfunction, premature coronary artery disease, gonadal dysfunction, postsplenectomy sepsis, and second malignancies. The second malignancies encountered include myelodysplasia and acute myeloid leukemia, non-Hodgkin's lymphomas, breast cancer, lung cancer, and melanoma. The major risk factor for hematologic malignancies is treatment with alkylating agents, while solid tumors are more likely to be seen with the use of radiation therapy. Patients cured of Hodgkin's disease seem to have greater fatigue, more psychosocial and sexual problems, and report a poorer quality of life than patients cured of acute leukemia.

Non-Hodgkin's Lymphomas The patient cured of a non-Hodgkin's lymphoma may be at increased risk of myelodysplasia and acute leukemia if high doses or prolonged alkylating agents were used. Chronic exposure to cyclophosphamide increases the risk of bladder cancer. Patients cured of lymphoma report a very good quality of life.

Acute Leukemia The late effects of anti-leukemic therapy include second malignancies (hematologic and solid tumors), neuropsychiatric difficulties, subnormal growth, thyroid abnormalities, and infertility.

Head and Neck Cancer Patients frequently have poor dentition, dry mouth, trismus, difficulty in eating, and poor nutrition. Those with nasopharyngeal cancer report the poorest long-term quality of life, possibly related to the volume of disease that is radiated.

Stem Cell Transplantation Cured patients are at risk of second cancers, especially if radiation therapy was part of the treatment. They are also subject to gonadal damage and infertility. Graft-versus-host disease is the leading factor contributing to the morbidity and mortality from allogeneic bone marrow transplantation, with an immune-mediated attack against the skin, liver, and gut epithelium. About half of patients report psychosexual problems.

Breast Cancer Patients treated with adjuvant chemotherapy and/or hormonal therapy for breast cancer are at risk for endometrial cancer from the use of tamoxifen. Those patients who have received chemotherapy may be at risk from doxorubicin or radiation-induced cardiomyopathy and acute leukemia. The development of premature ovarian failure from chemotherapy may cause hormone-deficient symptoms (hot flashes, decreased vaginal secretions, dyspareunia) and places women at risk for osteoporosis and cardiovascular deaths. Patients commonly report intrusive thoughts of cancer and psychological distress.

Testicular Cancer Depending on the modalities used for therapy, patients cured of testicular cancer can anticipate Raynaud's phenomena, renal and/or pulmonary damage from chemotherapy, and ejaculatory dysfunction from retroperitoneal lymph node dissection. Sexual dysfunction is reported by 15% of patients cured of testicular cancer.

Colon Cancer To date the major threat to patients with colorectal cancer treated with chemotherapy and or radiation therapy remains the risk of a second colorectal cancer. Quality of life is reported as high in long-term survivors.

Prostate Cancer Radical surgical treatment is often accompanied by impotence and about 10 to 15% develop some urine incontinence. Use of radiation therapy increases the risk of second cancers.

The challenge for the future is to integrate new chemotherapy and biologic agents and newer techniques of delivering radiation therapy in a fashion that increases cure rates and lowers the late effects of treatment. Additional populations at risk for late effects include those with cancers where therapy is becoming more effective, such as ovarian cancer, and cancers where chemotherapy and radiation therapy are used together in an organ-sparing approach, such as bladder cancer, anal cancer, and laryngeal cancer. Patients who have been cured of a cancer represent an important resource for cancer prevention studies.

(Bibliography omitted in Palm version)

Back to Table of Contents

SECTION 2 -DISORDERS OF HEMATOPOIESIS

104. HEMATOPOIESIS - *Peter J. Quesenberry*, *Gerald A. Colvin*

Hematopoiesis is the production of blood cells. It is a tightly regulated system exquisitely responsive to functional demands. The level of neutrophils, eosinophils, and basophils are maintained in discrete ranges with rapid adjustments when demands such as bacterial infection, parasitic infection, or allergic reaction are imposed. Similarly, lymphocytes, monocytes, platelets, and red cells, while maintained in normal ranges, respond rapidly to demands -- lymphocytes to immune challenge, monocytes to various infections, platelets to hemorrhage or inflammation, and red cells to tissue hypoxia from many causes. Derangements in marrow function can lead to an excess of white cells, such as leukemia or leukemoid reactions, or an inadequate number of cells, such as anemia, thrombocytopenia, or leukopenia. Kinetics of cytopenia induction after marrow injury with drugs, radiation, or infections reflect the life span of these cells in peripheral blood. The first lineage to drop are the neutrophils with a blood life span of 6 to 8 h, followed by platelets with a 10-day life span. Anemia develops over a longer time in the absence of blood loss, reflecting the 120-day life span of red blood cells. All of these cell types are produced by primitive cells termed *stem cells*, which are present in the bone marrow of adult mammals.

The production of all the cell types except lymphocytes is usually very efficient, and production is controlled largely by negative feedback. When demand for production of cells of a particular lineage increases or peripheral levels of the cells fall, stimulatory cytokines are released and generate new cells with a time delay of a few days, or the time required for maturation from stem cell precursors. By contrast, production of lymphocytes is highly inefficient. Each day many more cells are generated than are required in the periphery. Most lymphocytes are destroyed during development; this is due at least in part to the destruction of cells that express antigen receptors specific for self antigens.

HEMATOPOIETIC STEM CELLS

Hematopoietic stem cells are characterized by extensive proliferation and differentiation capacity, with the ability to self-renew on a population basis (Fig. 104-1). They also express a variety of cell-surface proteins and have the ability to rapidly "home" to bone marrow after intravenous injection. Human stem cells lack markers of lineage commitment (i.e., lineage-negative) and express c-Kit, c-mpl, and usually cluster of differentiation determinant-34 (CD34); a small subset of stem cells may be CD34-negative. Murine cells are also lineage-negative and express c-Kit, CD34, c-mpl, and Ly6A or Sca. The most primitive cells are characterized by low-level expression of a relatively large number of cytokine receptors and by relative exclusion (or pumping out) of the dyes rhodamine and Hoechst. These cells express a variety of adhesion proteins presumptively involved in marrow homing, including alpha$_4$, alpha$_5$, alpha$_6$, L-selectin, and platelet/endothelial cell adhesion molecule (PECAM) (Fig. 104-2). Another characteristic of the stem cell is a functional plasticity in response to cytokines as it transits the cell cycle (Fig. 104-3). Engraftment capacity is good in G_1 but virtually lost in late S and early G_2.

The stem cell is also a highly mobile cell with the capacity to evolve rapidly or involute pseudopodial extensions. The gold standard for defining the stem cell has been in vivo repopulation and long-term reconstitution of lethally irradiated mice. In vivo repopulation studies using unique radiation-induced chromosomal abnormalities or retroviral markers have shown that one or, at most, a few stem cells are capable of reconstituting the entire lymphohematopoietic system of a mouse; they also have defined classes of stem cells with short, medium, or long-term repopulating capacity. These are cell types that differ in the kinetics of hematopoietic reconstitution. Short-term cells repopulate in the first few weeks after transplantation but are not long-lasting; long-term cells account for long-lived reconstitution beginning a few months after reconstitution and lasting the entire life span; medium-term cells bridge the time between short- and long-term cells. When relatively small numbers of marked stem cells -- obtained by limiting dilution of sorted marrow cells -- are transplanted, lymphohematopoiesis may be clonal or oligoclonal, initially. Normal polyclonal lymphohematopoiesis derives from a relatively large number of clones. Both competitive marrow repopulation and mathematical studies support the model of polyclonal hematopoiesis. The most primitive long-term repopulating cells on activation with cytokines can rapidly alter phenotype and become short-term repopulating cells.

While stable multilineage chimerism has been documented in humans after clinical marrow transplantation, no assay system exists for human stem cell activity. A number of surrogate assays are used for the long-term, multilineage-repopulating cell in both humans and mice. These include the multifactor-responsive, high-proliferative potential colony-forming cells (HPP-CFC) and variations of stromal-based assays including the cobblestone-forming cell, long-term culture-initiating cell (LTC-IC) or LTC-IC-extended (LTC-IC-e). The adequacy of these assays is still the subject of debate. In addition, the NOD-SCID immunodeficient mouse has become a surrogate model for assaying human hematopoietic stem cells, although lineage skewing and variability of engraftment undermine its reliability.

LINEAGE PLASTICITY OF STEM CELLS

Tissue stem cells are capable of producing a wide variety of differentiated cell lineages, depending on intrinsic cell programming and the microenvironmental signals. Marrow cells may differentiate into mesenchymal, myocyte, endothelial, hematopoietic, and neural cells. Neural muscle and hepatic stem cells have been reported to give rise to hematopoiesis in transplanted mice. The regulation of stem cell plasticity and life span remains incompletely understood. However, once a particular set of transcription factors has been induced, either through an intrinsic program or from extracellular signals, reversibility is limited. The sequentially ordered activation of transcription factors leads to lineage commitment.

MICROENVIRONMENT

Nonhematopoietic tissues exert major influences on hematopoiesis, both short- and long-range. The nonhematopoietic tissues immediately abutting hematopoietic tissue have been termed the hematopoietic microenvironment, and the cells that comprise the environment influence hematopoiesis. For example, a surface location of adoptively transferred murine stem cells in the spleen of lethally irradiated mice favored

erythropoiesis, while an intrasplenic trabecular location was biased toward granulocyte production. In both human and murine species, various cell types have been identified in stroma, including hematopoietically derived macrophages and nonhematopoietic preadipocytic fibroblasts, endothelial cells, and vascular smooth muscle. This system appears capable of supporting the most primitive stem cells and controlling their proliferation and self-renewal. Most stem cells are resting under normal physiologic conditions but can be recruited into the cell cycle by demands of increased terminally differentiated hematopoietic cells. Cell-cell contact is critical in determining the microenvironment stimulus. Stem cells and primitive cells bind tightly to the stroma, while maturing precursors and terminally differentiated cells are nonadherent. Blocking interactions between stem cells and stromal cells with antibodies to vascular cell adhesion molecule (VCAM)-1 on stromal cells or its ligand, VLA-4, on stem cells block the interaction. Cytokine receptors binding to membrane-associated cytokines like stem cell factor, or to extracellular matrix-bound ligands, contribute other adhesive interactions.

PROGENITORS

Bone marrow stem cells can be induced to proliferate and differentiate into a wide variety of mature cell types in vitro in the presence of an appropriate colony-stimulating factor (CSF). Cells that give rise to mature colonies of granulocytes and macrophages are called granulocyte-macrophage colony-forming units (CFU-GM) (Fig. 104-4). The particular hematopoietic growth factor that stimulates the development of these colonies is called granulocyte-macrophage colony-stimulating factor (GM-CSF). Distinct culture conditions and supplemental growth factors, alone and in combination, are capable of producing a range of cell expansions from multilineage colonies that include lymphocytes to single-lineage clones. The different stem/progenitor clones are summarized in Table 104-1. Progenitor cells in general are found to have a higher proliferative rate and more lineage restriction than stem cells. They are also responsive to smaller numbers of cytokines. Thus, they are defined by expression of a limited variety of cytokine receptors.

The size of the colonies denotes the activity of cells at different stages of differentiation. Terminally acting cytokines produce smaller colonies called CFU (colony-forming units). When progenitors are stimulated with mixtures of early- and late-acting cytokines and are cultured for longer periods of time, the colonies are larger and multiple lineages are represented. Primitive multifactor-responsive erythroid colonies are termed burst-forming unit erythroid (BFU-E) while even more primitive colonies with great proliferative potential are termed HPP-CFC.

CYTOKINES

The lymphohematopoietic stem/progenitor populations and their progeny are largely defined by their cytokine responsiveness and cytokine receptor phenotype. Major efforts to define the regulators of granulocyte, erythroid, and platelet production have culminated in the definitions of a variety of glycoproteins. Acting through cell surface receptors at very low concentrations, these glycoproteins control the production of stem cells in vivo. Most prominent have been erythropoietin for red blood cells,GM-CSFfor granulocytes and macrophages, granulocyte-CSF (G-CSF) for granulocytes, and

thrombopoietin for platelets. In addition, macrophage-CSFor CSF-1, was defined as a primary regulator of macrophage-monocyte production and function. These cytokines exert prominent actions on specific cell lineages, but all exert actions on different cell lineages or on cells that have the potential to differentiate along more than one lineage.

In addition to the more lineage-restricted cytokines, a large number (perhaps up to 70) act broadly on multiple lineages and at multiple stages of lymphohematopoiesis. They exert effects on renewal, proliferation, survival, and differentiation; these effects may be stimulatory or inhibitory, and the cytokines usually show additive or synergistic effects with other cytokines. The cytokines also modulate intrinsic functions of early stem cells (migration and cell adherence) and promote the effector functions of their terminally differentiated progeny.G-CSFprimes neutrophils to undergo oxidative metabolism in response to formyl-methionyl-leucyl-phenylalanine (fMLF) and enhances cell migration, while interleukin (IL) 3 activates basophils, mast cells, and eosinophils.CSF-1 at low levels supports survival of murine marrow macrophages and at higher levels stimulates protein synthesis, cell division, and various macrophage functions, including antitumor activity, secretion of products of oxygen reduction, and plasminogen activator. CSF-1 also induces secretion of IL-1 from macrophages. Many of the hematopoietically active cytokines induce secretion of other cytokines, either inhibitory or stimulatory, creating multiple cytokine regulatory loops. Transforming growth factor b(TGF-b) is an inhibitory cytokine but also an autocrine factor supporting survival of pluripotent hematopoietic stem cells by blocking G1 to S phase transition. TGF-b conversely shows stimulatory effects on progenitors. Cytokines also modulate adhesion protein and integrin expression on multiple cell types. They exert their effects by interacting with surface-based receptors and initiating second-messenger cascades (see below).

The lymphohematopoietic cytokines can be broadly divided into colony-stimulating factors, erythropoietin, thrombopoietin, the interleukins, the inhibitory cytokines, chemokines that regulate cell migration and activation, and a variety of other hematopoietically active cytokines. A noninclusive overview of these cytokines emphasizing their primary, highlighted, or first-described action is presented in Tables 104-2,104-3and104-4. The general characteristics of cytokines are summarized inTable 104-5.

CYTOKINE RECEPTORS, SIGNAL TRANSDUCTION, AND TRANSCRIPTION FACTORS

Cytokines induce their effects through cell-surface membrane receptors. Several cytokine receptor families have been identified. The hematopoietic receptor family includesIL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-9,G-CSF,GM-CSF, and erythropoietin. Common characteristics of this family include four conserved cysteine residues and a WSXWS motif (X is a variable, nonconserved amino acid). Some also have immunoglobulin-like structures in their extracellular domains. Receptors frequently consist of multiple chains, and dimerization on cytokine binding is a usual feature of receptor biology. These receptors have no intrinsic signaling capacity and transmit signals by attaching to intracellular signaling molecules, such as the src family and the JAK family kinases. GM-CSF, IL-3, and IL-5 receptors have low-affinity alpha chains and a common high-affinity beta chain. The common beta chain may play a role in the competitive binding of these ligands.

Receptors for FLT-3 ligand, c-kit, platelet-derived growth factor (PDGF),CSF-1, and thrombopoietin constitute the tyrosine kinase receptor family. These receptors have conserved cysteines in the extracellular domain, with tyrosine kinase activity in the cytoplasmic domain, an immunoglobulin-like structure involved in ligand, and binding. Chemokine receptors are seven-transmembrane (serpin) G-protein linked receptors that signal cell activation and migration.

Cytokines typically cause receptor oligomerization on hematopoietic cells, followed by activation of intrinsic (receptor) or extrinsic tyrosine kinases, phosphorylation of the receptor and recruitment of Src-homology (SH2), and phospho-tyrosine binding (PT3) domain proteins to the receptor. Subsequent steps vary with different cytokines but essentially represent a series of phosphorylation-dephosphorylation events, with the final activation or nuclear translocation of a protein or protein complex that binds specific regions of DNA and initiates various genetic programs (i.e., acts as a transcription factor).

The complexity of these second-messenger signaling systems is illustrated by signaling through theGM-CSF,IL-3, and IL-5 receptors, which share a common beta chain. The beta chain does not have kinase activity but induces tyrosine phosphorylation of itself and a number of cytoplasmic proteins, including kinases, such as P1-3 kinase; adapters illustrated by Grb2; the insulin receptor-substrate 2 Cbl and Shc; guanine nucleotide exchange factors such as Vav; phosphastases such asSH2-domain protein tyrosine phosphatase-2 and SH2-containing inositol phosphatase; and transcription factors such as STAT 5.

Receptor phosphorylation is mediated by receptor-associated kinases, such as JAK2 (Janus family kinase 2, named Janus for the Roman god who guards the gates and looks in two directions; original Janus kinases were felt to have both tyrosine and serine kinase activity) and Src-family kinases. These sequential protein interactions lead to the evolution of proteins or protein complexes, termed *transcription factors*, that bind to specific regions of DNA to initiate genetic programs determining survival, proliferation, differentiation, and function.

As with second messengers, the transcription factor field is complex and evolving, but a number of transcription factors associated with specific stem cell levels or differentiation pathways have been described. Transcription factors that act at the earliest stem cell levels include c-myb, p45-NF-E2, GATA-2, AML-1 and tal-1/SCL, while Ikaros and PU-1 may act at the earliest lymphoid level. GATA-1 influences erythroid, mast cell, and megakaryocyte lineages, while FOG (friend of GATA-1) acts in concert with GATA-1. PU-1 appears to influence granulocyte and monocyte differentiation, P45-NF-E2 affects megakaryocyte lineages, and PAX-5 B lymphoid development. These transcription factors usually act in complexes with specific conformations binding to particular DNA sequences.

MIGRATION HOMING AND ADHESION PROTEINS

The process of stem cell homing to the marrow is complex and involves a number of adhesion proteins. Very late antigen (VLA) 4, VLA-5, VLA-6,PECAM, P- and E-selectin,

CD44, CXCR, and a receptor for ligand-bearing galactosyl and mannosyl residues have been shown to be expressed by hematopoietic stem/progenitor cells and implicated in marrow homing. The integrins α_4 and/or α_5 are expressed on immature blasts, erythroid progenitors, monocytes, and CD34+ cells; in general, expression of α_4 appears to decrease with maturation. Hematopoietic cells also bind differentially to different extracellular matrix components: erythroid cells to fibronectin, CFU-GM and BFU-E to collagen.

Antibody to VLA-4 given in vivo causes mobilization of hematopoietic progenitors in normal or cytokine-treated primates and/or mice. Stem cell mobilization by cytokines involves down regulation of adhesion protein expression on hematopoietic stem cells. The cell-cycle related fluctuations in engraftment appear to be based on alterations on different surface adhesion proteins. The stem cells are highly motile and move in a direction of cytokine or chemokine gradients with steel factor and stromal-derived factor 1 (SDF-1) being active. Adhesion proteins act not only for motility/adhesion, but also serve a regulatory role that is similar in some cases to traditional cytokines.

PHYSIOLOGY OF HEMATOPOIESIS AND SOURCES OF CYTOKINES

Erythropoietin is produced largely by the kidney in response to tissue hypoxia. The regulation of granulocyte and monocyte production is more complex, but appears to be in response to various infectious or noxious agents, such as gram-negative bacteria, the endotoxin in the cell wall of these bacteria, and antigen stimulation. All of these interact with peripheral tissue cells to generate a variety of cytokine messages, resulting in increase production in specific cell types. Parasitic infections appear to elicit IL-5, which modulates the eosinophilia and mast cell lineages. Viral infections have specific effects on lymphocyte classes; typically bacterial infections stimulate granulocyte production. Tuberculosis or other mycobacterial infections predominantly induce increased monocyte production. All of these biologic affects appear to be mediated by the selective evolution of cytokine complexes from tissue endothelial cells, fibroblasts, macrophages, and lymphocytes. Most cells produce a large variety of cytokines, but the key is the relative levels, combinations, and timing of the production of these cytokines (Fig. 104-5).

HEMATOPOIETIC STEM CELL AND CYTOKINE DISEASES

The classic stem cell disease is *chronic myeloid leukemia*. Here a specific genetic translocation between chromosomes 9 and 22 at the stem cell level leads to excess production of granulocytes, monocytes, basophils, frequently platelets, and less frequently red cells. Other lymphohematopoietic clonal stem cell diseases include polycythemia vera, myelofibrosis with myeloid metaplasia, paroxysmal nocturnal hemoglobinuria, and acute myeloid leukemia. Out of the scope of this chapter, but relevant to these discussions, is the fact that many lymphoid neoplasms are clonal diseases at early stages of development, but probably not at the mature stage suggested by the tumor cell-surface phenotype. The vast majority of peripheral B cell and T cell malignancies have genetic lesions associated with receptor gene rearrangements, which occur early in lymphoid cell development. Aplastic anemia appears to be a disease characterized by a defective number of hematopoietic stem cells. Cyclic hematopoiesis is another disease of hematopoietic stem cells. In gray collie

dogs with this disorder, levels of platelets, reticulocytes, monocytes, and granulocytes cycle. This disease can be cured or transmitted by marrow transplantation. The human disease, cyclical neutropenia -- or cyclic hematopoiesis in which blood cells oscillate with a 21-day period -- is caused by missense and splicing mutations in the gene encoding neutrophil elastase, thus implicating this inflammatory chymotryptic serine protease in the oscillatory timing of hematopoiesis. The stem cell diseases are summarized inTable 104-6.

A number of cytokine disorders or diseases have now been defined. The best characterized is the anemia of renal failure, an erythropoietin deficiency state that can be corrected by the administration of erythropoietin. Various tumors, particularly lung cancer, increase peripheral granulocyte counts secondary to the production ofG-CSF. The IL-6 family of cytokines appears to be prominently involved in a number of inflammatory states, causes the systemic symptoms associated with Castleman's disease and atrial myxoma, and may be an etiologic factor in multiple myeloma. IL-6 may also be a major cause of symptoms in various lymphomas. Abnormalities of the c-Kit receptor may underlie a number of mast cell diseases in humans; IL-5 production is the proximate cause of a number of eosinophilic states. A deficiency of IL-1 is a feature of aplastic anemia. Mutations in the G-CSF receptor in chronic congenital neutropenia (Kostmann's syndrome) may be a causative factor in the evolution of acute myeloid leukemia in some of these patients.

THERAPEUTIC IMPLICATIONS OF STEM CELLS AND CYTOKINES

Stem Cells Stem cell transplantation was first established as an effective therapy for relapsed acute myeloid leukemia and aplastic anemia. It is now a mainstay of therapy for virtually all leukemias and some relapsed lymphomas. Application of this treatment to a number of solid tumors has been disappointing. Major expectations with regard to its potential in breast cancer have not yet been fulfilled, although it appears to be effective in relapsed testicular cancer. The rationale is the use of very high doses of drugs or radiation designed to kill all tumor cells, but at levels where marrow toxicity would be lethal. Marrow damage is the dose-limiting toxicity for many chemotherapeutic agents. If marrow function is replaced by transplant, it might be possible to increase the dose of chemotherapy substantially before another organ toxicity becomes dose-limiting.

The strategy is somewhat different in intrinsic marrow diseases, such as aplastic anemia, where marrow function is restored without the need for killing tumor cells, or in genetic marrow diseases such as thalassemia and sickle cell anemia, where replacement of the abnormal marrow with normal marrow corrects the disease state. Aggressive autoimmune diseases are also being treated with marrow replacement. Stem cells can come from a related or unrelated allogeneic source or directly from the patient. The sources of the stem cells also vary. Initially, marrow aspirate was the predominant source, but now apheresed peripheral blood stem cells are the most utilized. In addition, umbilical vein cord blood, especially in pediatric patients, appears effective. Numbers of cells are sometimes limiting for adult recipients (Chap 115).

Cytokines Demonstration that hematopoietic cytokines could modulate red blood cell and white cell production in humans has been useful in some clinical settings.

Erythropoietin treatment improves hematocrit and quality of life in patients with chronic renal failure. Erythropoietin has been tried in myelodysplastic syndromes (MDS), a group of clonal stem cell disorders. Meta-analysis shows an overall response rate of only 13%; the actual overall clinical benefit was exceedingly small. In addition, the best results were seen in patients receiving daily injections. Prohibitive costs and often poor response limits use to those patients with serum EPO levels lower than 500 mU/L, and less than 5% myeloblasts. The utilization of erythropoietin in other settings than renal failure remains controversial and its use may relate more to effective marketing than to the science.

These concerns are multiplied for the use of the myeloid growth factors,G-CSF, andGM-CSF. These agents elevate neutrophil and monocyte counts, and under very selective conditions they can result in a reduced toxicity of various chemotherapeutic regimens (Table 104-7). Unfortunately, they save about the same amount of money in hospitalizations for febrile neutropenia as they cost, and their use has not increased survival rate. Virtually all of the G-CSF and GM-CSF trials in cancer patients have been flawed by design; they involve escalation of drugs to toxic levels and reversal of toxicity without addressing the question of whether the patient's survival is affected by the treatment. GM-CSF and G-CSF are grossly overutilized; they are often used in settings where their efficacy has not been shown (e.g., patients with a low probability of neutropenia). Their use should still be considered experimental, and they should continue to be studied in a protocol setting but not used routinely. G-CSF is useful in mobilization of stem cells, and it is also effective in treatment of various chronic neutropenias, in particular cyclic neutropenia and Kostmann's syndrome. G-CSF may be involved in the evolution to acute myeloid leukemia in some patients, but overall it appears to be an effective intervention in these seriously ill patients. G-CSF may also aid in healing of diabetic skin ulcers.

IL-11 and thrombopoietin can elevate platelet counts in experimental animals, but their place in clinical practice is unclear. IL-11 has been approved for use in chemotherapy-induced thrombocytopenia, but its effects are small. Use of pegylated recombinant human megakaryocyte growth and development factor -- the truncated version of thrombopoietin -- has resulted in production of neutralizing antithrombopoietin antibodies and thrombocytopenia. Recombinant human thrombopoietin does not commonly elicit neutralizing antithrombopoietin antibodies. Clinical benefit (or cost effectiveness) has not yet been shown with thrombopoietin. Surrogate values of platelet counts or number of platelet transfusions are not valid criteria for clinical benefit. Thrombopoietin may eventually find a role as an expander of early stem cells in vitro. Active research in this important area continues.

Gene Therapy Hematopoietic stem cells provide an ideal vehicle for various gene therapy approaches. These cells can be easily induced into the cell cycle for retroviral integration. Long-term expression of introduced genes is currently being obtained in animal models. Initial clinical application has been disappointing, but success has been obtained in Gaucher's disease, suggesting that gene therapy will eventually become a successful approach to a number of hematopoietic diseases.

ACKNOWLEDGEMENT

Dr. Francis Ruscetti and Dr. Jonathan Keller contributed this chapter in the 14th

edition and portions of their chapter have been retained here.

(Bibliography omitted in Palm version)

Back to Table of Contents

105. IRON DEFICIENCY AND OTHER HYPOPROLIFERATIVE ANEMIAS - *John W. Adamson*

Anemias associated with normocytic and normochromic red cells and an inappropriately low reticulocyte response (reticulocyte index <2.5) are *hypoproliferative anemias*. This category includes early iron deficiency (before hypochromic microcytic red cells develop), acute and chronic inflammation (including many malignancies), renal disease, hypometabolic states such as protein malnutrition and endocrine deficiencies, and anemias from marrow damage. Marrow damage states are discussed in Chap. 109. Hypoproliferative anemias are the most common anemias and anemia associated with acute and chronic inflammation is the most common of these. The anemia of acute and chronic inflammation, like iron deficiency, is related in part to abnormal iron metabolism. The anemias associated with renal disease, inflammation, cancer, and hypometabolic states are characterized by an abnormal erythropoietin response to anemia.

IRON METABOLISM

Iron is a critical element in the function of all cells, although the amount of iron required by individual tissues varies during development. At the same time, the body must protect itself from free iron, which is highly toxic in that it participates in chemical reactions that generate free radicals such as singlet O_2 or OH^-. Consequently, elaborate mechanisms have evolved that allow iron to be made available for critical physiologic functions while at the same time conserving this element and handling it in such a way that toxicity is avoided.

The major role of iron in mammals is to carry O_2 as part of the heme protein that, in turn, is part of hemoglobin. O_2 also is bound by a heme protein in muscle, myoglobin. Iron also is a critical element in iron-containing enzymes, including the cytochrome system in mitochondria. Iron distribution in the body is shown in Table 105-1. Without iron, cells lose their capacity for electron transport and energy metabolism; in erythroid cells hemoglobin synthesis is impaired, resulting in anemia and reduced O_2 delivery to tissue.

THE IRON CYCLE IN HUMANS

Figure 105-1 outlines the major pathways of internal iron exchange in humans. Iron absorbed from the diet or released from stores circulates in the plasma bound to *transferrin*, the iron transport protein. Transferrin is a bilobed glycoprotein with two iron binding sites. Transferrin that carries iron exists in two forms -- *monoferric* (one iron atom) or *diferric* (two iron atoms). The turnover (half-clearance time) of transferrin-bound iron is very rapid -- typically 60 to 90 min. Because the overwhelming majority of iron transported by transferrin is delivered to the erythroid marrow, the clearance time of transferrin-bound iron from the circulation is affected most by the plasma iron level and the activity of the erythroid marrow. When erythropoiesis is markedly stimulated, the pool of erythroid cells requiring iron increases and the clearance time of iron from the circulation decreases. The half-clearance time of iron in the presence of iron deficiency is as short as 10-15 min; this value reflects the limits of iron delivery as a function of the cardiac output going to the bone marrow. With suppression of the erythroid marrow, the plasma iron level typically is increased and the half-clearance time is prolonged to as much as several hours. Normally, the iron bound

to transferrin turns over 10 to 20 times per day. Assuming a normal plasma iron level of 80 to 100 ug/dL, the amount of iron passing through the transferrin pool is 20 to 24 mg/d.

The iron-transferrin complex circulates in the plasma until the iron-carrying transferrin interacts with specific *transferrin receptors* on the surface of marrow erythroid cells. Diferric transferrin has the highest affinity for transferrin receptors; apotransferrin (transferrin not carrying iron) has very little affinity. While transferrin receptors are found on cells in many tissues within the body -- and all cells at some time during development will display transferrin receptors -- the cell having the greatest number of receptors (300,000 to 400,000/cell) is the developing erythroblast.

Once the iron-bearing transferrin interacts with its receptor, the iron-transferrin-receptor complex is internalized via clathrin-coated pits and transported to an acidic endosome, where the iron is released at the low pH. The iron is then made available for heme synthesis while the transferrin-receptor complex is recycled to the surface of the cell, where the bulk of the transferrin is released back into the circulation and the transferrin receptor reanchors into the cell membrane. At this point a certain amount of the transferrin receptor protein may be released into circulation. Within the erythroid cell, iron that is in excess of the amount needed for hemoglobin synthesis binds to a storage protein, *apoferritin*, forming *ferritin*. This mechanism of iron exchange also takes place in other cells of the body expressing transferrin receptors, especially liver parenchymal cells where the iron can be incorporated into heme-containing enzymes or stored. The iron incorporated into hemoglobin subsequently enters the circulation as new red cells are released from the bone marrow. The iron is then part of the red cell mass and will not become available for reutilization until the red cell dies.

In a normal individual, the average red cell life span is 120 days. Thus, 0.8 to 1.0% of red cells turn over each day. At the end of its life span, the red cell is recognized as senescent by the cells of the *reticuloendothelial (RE) system*, and the cell undergoes phagocytosis. Once within the RE cell, the hemoglobin from the ingested red cell is broken down, the globin and other proteins are returned to the amino acid pool, and the iron is shuttled back to the surface of the RE cell, where it is presented to circulating transferrin. The "harvesting" of iron from senescent red cells is both efficient and rapid, with newly recycled iron appearing in the circulation within 10 min of ingestion of the red cell. It is the efficient and highly conserved recycling of iron from senescent red cells that supports steady state (and even mildly accelerated) erythropoiesis.

Since each milliliter of red cells contains 1 mg of elemental iron, the amount of iron needed to replace those red cells lost through senescence amounts to 16 to 20 mg/day (assuming an adult with a red cell mass of 2 L). Any additional iron required for daily red cell production comes from the diet. Normally, an adult male will need to absorb at least 1 mg of elemental iron daily to meet needs, while females in the childbearing years will need to absorb an average of 1.4 mg/d. However, to achieve a maximum proliferative erythroid marrow response to anemia, additional iron must be available. With markedly stimulated erythropoiesis, demands for iron are increased by as much as six- to eightfold. With hemolytic anemias, the rate of red cell destruction is increased, but the iron recovered from the red cells is efficiently reutilized for hemoglobin synthesis. In contrast, with blood loss anemia the rate of red cell production is limited by the amount

of iron that can be mobilized from ferritin and hemosiderin stores. Typically, the rate of mobilization under these circumstances will not support red cell production more than 2.5 to 3 times normal. If the delivery of iron to the stimulated marrow is suboptimal, the marrow's proliferative response is blunted and normal hemoglobin synthesis is impaired. The result is a hypoproliferative marrow accompanied by microcytic, hypochromic anemia.

While blood loss or hemolysis places a demand for iron to be supplied to the erythroid marrow, other conditions such as inflammation interfere with iron release from stores and can result in a rapid decrease in the serum iron (see below).

NUTRITIONAL IRON BALANCE

The balance of iron metabolism in the organism is tightly controlled and designed to conserve iron for reutilization. There is no excretory pathway for iron, and the only mechanisms by which iron is lost from the body are blood loss (via gastrointestinal bleeding, menses, or other forms of bleeding) and the loss of epidermal cells from the skin and gut. Normally, the only route by which iron comes into the body is via absorption from food (dietary iron intake) or from medicinal iron taken orally. Iron may also enter the body through red cell transfusions or injection of iron complexes. The margin between the amount of iron available for absorption and the requirement for iron in growing infants and the adult female is narrow. The narrowness of this margin accounts for the great prevalence of iron deficiency worldwide -- currently estimated at one-half billion people.

External iron exchange -- the amount of iron required from the diet to replace losses -- averages about 10% of body iron content a year in the male and 15% in women of childbearing age, equivalent to 1.0 and 1.4 mg of elemental iron daily, respectively. Dietary iron content is closely related to total caloric intake (approximately 6 mg of elemental iron per 1000 calories). Iron bioavailability is affected by the nature of the foodstuff with heme iron (e.g., red meat) being most readily absorbed. In the United States, the average iron intake in an adult male is 15 mg/d with 6% absorption; for the average female, the daily intake is 11 mg/d with 12% absorption. An individual with iron deficiency can increase iron absorption to about 20% of the iron present in a meat-containing diet but only 5 to 10% of the iron in a vegetarian diet. As a result, nearly one-third of the female population in the United States has virtually no iron stores. Vegetarians are at an additional disadvantage because certain foodstuffs that include phytates and phosphates reduce iron absorption by about 50%. When ionizable iron salts are given together with food, the amount of iron absorbed is reduced. This is particularly true with iron in the ferric state. When the percentage of iron absorbed from individual food items is compared with the percentage for an equivalent amount of ferrous salt, iron in vegetables is only about one-twentieth as available, egg iron one-eighth, liver iron one-half, and heme iron one-half to two-thirds. Therefore, liver and heme iron are absorbed nearly as well as iron salt added to food, while the iron in vegetables and eggs is much less available.

Infants, children, and adolescents may be unable to maintain normal iron balance because of the demands of body growth and lower dietary intake of iron. In pregnancy during the last two trimesters, daily iron requirements increase to 5 to 6 mg. That is the

reason why iron supplements are almost universally prescribed for pregnant women in developed countries. Enthusiasm for supplementing foods such as bread and cereals with iron has waned in the face of concerns that the very prevalent hemochromatosis gene would result in an unacceptable risk of iron overload.

Iron absorption takes place largely in the proximal small intestine and is a carefully regulated process. For absorption, iron must be taken up by the luminal cell. That process is facilitated by the acidic contents of the stomach, which maintains the iron in solution. At the brush border of the absorptive cell, the ferric iron is converted to the ferrous form by a ferrireductase. Transport across the membrane is accomplished by divalent metal transporter 1 (DMT 1, also known as Nramp 2 or DCT 1). DMT 1 is a general cation transporter. Once iron is inside the gut cell, the iron may be stored as ferritin or transported through the cell to be released at the basolateral surface to plasma transferrin. It is likely another transporter acts here in concert with hephaestin, another ferroxidase. Hephaestin is similar to ceruloplasmin, the copper-carrying protein.

Iron absorption is influenced by a number of physiologic states. Erythroid hyperplasia, for example, stimulates iron absorption, even in the face of normal or increased iron stores. Patients with anemias associated with high levels of ineffective erythropoiesis absorb excess amounts of dietary iron. Over time, this may lead to iron overload and tissue damage. In iron deficiency iron is much more efficiently absorbed from a given diet while the contrary is true in the presence of iron overload. This is possibly mediated through signals that become fixed before the jejunal crypt cell migrates up the villus to become an absorptive cell. The normal individual can reduce iron absorption in situations of excessive intake or medicinal iron intake; however, while the percent of iron absorbed goes down, the absolute amount goes up. This accounts for the acute iron toxicity occasionally seen when children ingest large numbers of iron tablets. Under these circumstances, the amount of iron absorbed exceeds the transferrin binding capacity of the plasma, resulting in free iron that affects critical organs such as cardiac muscle cells.

IRON DEFICIENCY ANEMIA

STAGES OF IRON DEFICIENCY

Iron deficiency anemia is the condition in which there is anemia and clear evidence of iron deficiency. However, it is worthwhile to consider the steps by which iron deficiency occurs (Fig. 105-2). These can be divided into three stages. The first stage is *negative iron balance*, in which the demands for (or losses of) iron exceed the body's ability to absorb iron from the diet. This stage can result from a number of physiologic mechanisms including blood loss, pregnancy (in which the demands for red cell production by the fetus outstrip the mother's ability to provide iron), rapid growth spurts in the adolescent, or inadequate dietary iron intake. Most commonly, the growth needs of the fetus or rapidly growing child exceed the individual's ability to absorb the iron necessary for hemoglobin synthesis from the diet. Blood loss in excess of 10 to 20 mL of red cells per day is greater than the amount of iron that the gut can absorb from a normal diet. Under these circumstances the iron deficit must be made up by mobilization of iron fromREstorage sites. During this period measurements of iron stores -- such as the serum ferritin level or the appearance of stainable iron on bone marrow aspirations

-- will decrease. As long as iron stores are present and can be mobilized, the serum iron, total iron-binding capacity (TIBC), and red cell protoporphyrin levels remain within normal limits. At this stage, red cell morphology and indices are normal.

When iron stores become depleted, the serum iron begins to fall. Gradually, theTIBCincreases, as do red cell protoporphyrin levels. By definition, marrow iron stores are absent when the serum ferritin level <15 ug/L. As long as the serum iron remains within the normal range, hemoglobin synthesis is unaffected despite the dwindling iron stores. Once the transferrin saturation falls to 15 to 20%, hemoglobin synthesis becomes impaired. This is a period of *iron-deficient erythropoiesis*. Careful evaluation of the peripheral blood smear reveals the first appearance of microcytic cells, and if the laboratory technology is available, one finds hypochromic reticulocytes in circulation. Gradually, the hemoglobin and hematocrit begin to fall, reflecting *iron deficiency anemia*. The transferrin saturation at this point is 10 to 15%.

When moderate anemia is present (hemoglobin 10-13 g/dL), the bone marrow remains hypoproliferative. With more severe anemia (hemoglobin 7-8 g/dL), hypochromia and microcytosis become more prominent, misshapen red cells (poikilocytes) appear on the blood smear as cigar or pencil-shaped forms and target cells, and the erythroid marrow becomes increasingly ineffective. Consequently, with severe prolonged iron deficiency anemia, erythroid hyperplasia of the marrow develops rather than hypoproliferation.

CAUSES OF IRON DEFICIENCY

Conditions that increase demand for iron, increase iron loss, or decrease iron intake, absorption, or use can produce iron deficiency (Table 105-2).

CLINICAL PRESENTATION OF IRON DEFICIENCY

Certain clinical conditions carry an increased likelihood of iron deficiency. Pregnancy, adolescence, periods of rapid growth, and an intermittent history of blood loss of any kind should alert the clinician to possible iron deficiency. A cardinal rule is that the appearance of iron deficiency in an adult male means gastrointestinal blood loss until proven otherwise. Signs related to iron deficiency depend upon the severity and chronicity of the anemia in addition to the usual signs of anemia -- fatigue, pallor, and reduced exercise capacity. *Cheilosis* (fissures at the corners of the mouth) and *koilonychia* (spooning of the fingernails) are signs of advanced tissue iron deficiency. The diagnosis of iron deficiency is typically based on laboratory results.

LABORATORY IRON STUDIES

Serum Iron and Total Iron-Binding Capacity The serum iron level represents the amount of circulating iron bound to transferrin. TheTIBC is an indirect measure of the circulating transferrin. The normal range for the serum iron is 50 to 150 ug/dL; the normal range for TIBC is 300 to 360 ug/dL. Transferrin saturation, which is normally 25 to 50%, is obtained by the following formula: serum iron ´ 100¸ TIBC. Iron deficiency states are associated with saturation levels below 18%. In evaluating the serum iron, the clinician should be aware that there is a diurnal variation in the value. A transferrin saturation rate of >50% indicates that a disproportionate amount of the iron bound to

transferrin is being delivered to nonerythroid tissues. If this condition persists for an extended time, tissue iron overload may occur.

Serum Ferritin Free iron is toxic to cells, and the body has established an elaborate set of protective mechanisms to bind iron in various tissue compartments. Within cells, iron is stored complexed to protein as ferritin or hemosiderin. Apoferritin binds to free ferrous iron and stores it in the ferric state. As ferritin accumulates within cells of theREsystem, protein aggregates are formed as hemosiderin. Iron in ferritin or hemosiderin can be extracted for release by the RE cells although hemosiderin is less readily available. Under steady state conditions, the serum ferritin level correlates with total body iron stores; thus, the serum ferritin level is the most convenient laboratory test to estimate iron stores. The normal value for ferritin varies according to the age and gender of the individual (Fig. 105-3). Adult males have serum ferritin values averaging about 100 ug/L, while adult females have levels averaging 30 ug/L. As iron stores are depleted, the serum ferritin falls to<15 ug/L. Such levels are virtually always diagnostic of absent body iron stores.

Evaluation of Bone Marrow Iron Stores AlthoughRE cell iron stores can also be estimated from the iron stain of a bone marrow aspirate or biopsy, the measurement of serum ferritin has largely supplanted bone marrow aspirates for determination of storage iron (Table 105-3). The serum ferritin level is a better indicator of iron overload than the marrow iron stain. However, in addition to storage iron the marrow iron stain provides information about the effective delivery of iron to developing erythroblasts. Normally, 40 to 60% of developing erythroblasts -- called *sideroblasts* -- will have visible ferritin granules in their cytoplasm. This represents iron in excess of that needed for hemoglobin synthesis. In states in which release of iron from storage sites is blocked, RE iron will be detectable, and there will be few or no sideroblasts. In the myelodysplastic syndromes, mitochondrial dysfunction occurs, and accumulation of iron in mitochondria appears in a necklace fashion around the nucleus of the erythroblast. Such cells are referred to as *ringed sideroblasts*.

Red Cell Protoporphyrin Levels Protoporphyrin is an intermediate in the pathway to heme synthesis. Under conditions in which heme synthesis is impaired, protoporphyrin accumulates within the red cell. This can reflect an inadequate iron supply to erythroid precursors to support hemoglobin synthesis. Normal values are less than 30 ug/dL of red cells. In iron deficiency, values in excess of 100 ug/dL are seen. The most common causes of increased red cell protoporphyrin levels are absolute or relative iron deficiency and lead poisoning.

Serum Levels of Transferrin Receptor Protein Because erythroid cells have the highest numbers of transferrin receptors on their surface of any cell in the body, and because transferrin receptor protein (TRP) is released by cells into the circulation, serum levels of TRP reflect the total erythroid marrow mass. Another condition in which TRP levels are elevated is absolute iron deficiency. Normal values are 4 to 9 ug/L determined by immunoassay. This laboratory test is becoming increasingly available and has been proposed to measure the serial expansion of the erythroid marrow in response to recombinant erythropoietin therapy.

DIFFERENTIAL DIAGNOSIS

Other than iron deficiency, only three conditions need to be considered in the differential diagnosis of a hypochromic microcytic anemia (Table 105-4). The first is inherited defects in globin chain synthesis: the thalassemias. These are differentiated from iron deficiency most readily by serum iron values, since it is characteristic to have at least normal -- if not increased -- serum iron levels and transferrin saturation with the thalassemias.

The second condition is chronic inflammatory disease with inadequate iron supply to the erythroid marrow. The distinction between true iron deficiency anemia and the anemia associated with chronic inflammatory states is among the most common diagnostic problems encountered by clinicians (see below). Usually the anemia of chronic disease is normocytic and normochromic. Again, the iron values usually make the differential diagnosis clear, as the ferritin level is normal or increased and theTIBC is typically below normal.

Finally, the myelodysplastic syndromes comprise the third condition. Some patients with myelodysplasia have impaired hemoglobin synthesis with mitochondrial dysfunction resulting in impaired iron incorporation into heme. The iron values again reveal normal stores and more than an adequate supply to the marrow, despite the microcytosis and hypochromia.

TREATMENT

The severity and cause of iron deficiency anemia will determine the appropriate approach to treatment. As an example, symptomatic elderly patients with severe iron deficiency anemia and cardiovascular instability may require red cell transfusions. Younger individuals who have compensated for their anemia can be treated more conservatively with iron replacement. The foremost issue for the latter patient is the precise identification of the cause of the iron deficiency.

For the majority of cases of iron deficiency (pregnant women, growing children and adolescents, patients with infrequent episodes of bleeding, and those with inadequate dietary intake of iron), oral iron therapy will suffice. For patients with unusual blood loss or malabsorption, specific diagnostic tests and appropriate therapy take priority. Once the diagnosis of iron deficiency anemia and its cause is made, and a therapeutic approach is charted, there are three major approaches.

Red Cell Transfusion Transfusion therapy is reserved for those individuals who have symptoms of anemia, cardiovascular instability, and continued and excessive blood loss from whatever source, and those who require immediate intervention. The management of these patients is less related to the iron deficiency than it is to the consequences of the severe anemia. Not only do transfusions correct the anemia acutely, but the transfused red cells provide a source of iron for reutilization, assuming they are not lost through continued bleeding. Transfusion therapy will stabilize the patient while other options are reviewed.

Oral Iron Therapy In the patient with established iron deficiency anemia who is asymptomatic, treatment with oral iron is usually adequate. Multiple preparations are

available ranging from simple iron salts to complex iron compounds designed for sustained release throughout the small intestine (Table 105-5). While the various preparations contain different amounts of iron, they are generally all absorbed well and are effective in treatment. Some come with other compounds designed to enhance iron absorption, such as citric acid. It is not clear whether the benefits of such compounds justify their costs. Typically, for iron replacement therapy up to 300 mg of elemental iron per day is given, usually as three or four iron tablets (each containing 50 to 65 mg elemental iron) given over the course of the day. Ideally, oral iron preparations should be taken on an empty stomach, since foods may inhibit iron absorption. Some patients with gastric disease or prior gastric surgery require special treatment with iron solutions, since the retention capacity of the stomach may be reduced. The retention capacity is necessary for dissolving the shell of the iron tablet before the release of iron. A dose of 200 to 300 mg of elemental iron per day should result in the absorption of up to 50 mg of iron per day. This supports a red cell production level of two to three times normal in an individual with a normally functioning marrow and appropriate erythropoietin stimulus. However, as the hemoglobin level rises, erythropoietin stimulation decreases, and the amount of iron absorbed is reduced. The goal of therapy in individuals with iron deficiency anemia is not only to repair the anemia, but also to provide stores of at least 1/2 to 1 g of iron. Sustained treatment for a period of 6 to 12 months after correction of the anemia will be necessary to achieve this.

Of the complications of oral iron therapy, gastrointestinal distress is the most prominent and is seen in 15 to 20% of patients. For these patients, abdominal pain, nausea, vomiting, or constipation often lead to noncompliance. Although small doses of iron or iron preparations with delayed release may help somewhat, the gastrointestinal side effects are a major impediment to the effective treatment of a number of patients.

The response to iron therapy varies, depending upon the erythropoietin stimulus and the rate of absorption. Typically, the reticulocyte count should begin to increase within 4 to 7 days after initiation of therapy and peak at 1 1/2 weeks. The absence of a response may be due to poor adsorption, noncompliance (which is common), or a confounding diagnosis. If iron deficiency persists, it may be necessary to switch to parenteral iron therapy.

Parenteral Iron Therapy Intramuscular or intravenous iron can be given to patients who are unable to tolerate oral iron, whose needs are relatively acute, or who need iron on an ongoing basis, usually due to persistent gastrointestinal blood loss. Currently, the intravenous route is used routinely. Parenteral iron use has been rising rapidly in the last several years with the recognition that recombinant erythropoietin therapy induces a large demand for iron -- a demand that frequently cannot be met through the physiologic release of iron from RE sources. Concern has been raised about the safety of parenteral iron -- particularly iron dextran. The serious adverse reaction rate to intravenous iron dextran is 0.7%. Fortunately, newer iron complexes are becoming available in the United States that are likely to have an even lower rate of adverse effects. The most recently approved preparation is intravenous iron gluconate (Ferrlecit).

There are two approaches to the use of parenteral iron: one is to administer the total dose of iron required to correct the hemoglobin deficit and provide the patient with at least 500 mg of iron stores; the second is to give repeated small doses of parenteral

iron over a protracted period. The latter approach is common in dialysis centers, where it is not unusual for 100 mg of elemental iron to be given weekly for 10 weeks to augment the erythropoietic response to recombinant erythropoietin therapy. The amount of iron needed by an individual patient is calculated by the following formula: body weight (kg) ´ 2.3´ (15 - patient's hemoglobin, g/dL) + 500 or 1000 mg (for stores).

In administering intravenous iron, anaphylaxis is always a concern. Anaphylaxis is less common with the newer preparations. The factors that have correlated with a serious anaphylactic-like reaction include a history of multiple allergies or a prior allergic reaction to dextran (in the case of iron dextran). Generalized symptoms appearing several days after the infusion of a large dose of iron can include arthralgias, skin rash, and low-grade fever. This may be dose-related, but it does not preclude the further use of parenteral iron in the patient. To date, patients with sensitivity to iron dextran have been safely treated with iron gluconate. If a large dose of iron dextran is to be given (>100 mg) the iron preparation should be diluted in 5% dextrose in water or 0.9% NaCl solution. The iron solution can then be infused over a 60 to 90 min period (for larger doses) or at a rate convenient for the attending nurse or physician. While a test dose (25 mg) of parenteral iron is recommended, in reality a slow infusion of a larger dose of parenteral iron solution will afford the same kind of early warning as a separately injected test dose. Early in the infusion of iron, if chest pain, wheezing, a fall in blood pressure, or other systemic manifestations occur, the infusion of iron -- whether as a large solution or a test dose -- should be interrupted immediately.

OTHER HYPOPROLIFERATIVE ANEMIAS

In addition to mild to moderate iron deficiency anemia, the hypoproliferative anemias can be divided into four categories: (1) chronic inflammation/infection; (2) renal disease; (3) endocrine and nutritional deficiencies (hypometabolic states); and (4) marrow damage (Chap. 109). With chronic inflammation, renal disease, or hypometabolism, endogenous erythropoietin production is inadequate for the degree of anemia observed. For the anemia of chronic inflammation (anemia of chronic disease), the erythroid marrow also responds inadequately to stimulation in part due to defects in iron reutilization. As a result of the lack of adequate erythropoietin stimulation, an examination of the peripheral blood smear will disclose only an occasional polychromatophilic (shift) reticulocyte. In the cases of iron deficiency or marrow damage, appropriate elevations in endogenous erythropoietin levels are typically found, and "shift" reticulocytes will be present on the blood smear.

ANEMIA OF ACUTE AND CHRONIC INFLAMMATION/INFECTION (THE ANEMIA OF CHRONIC DISEASE)

The anemia of chronic disease -- which encompasses inflammation, infection, tissue injury, and conditions associated with the release of proinflammatory cytokines (such as cancer) -- is one of the most common forms of anemia seen clinically and is probably the most important in the differential diagnosis of iron deficiency, since many of the features of the anemia are brought about by inadequate iron delivery to the marrow, despite the presence of normal or increased iron stores. This is reflected by a low serum iron, increased red cell protoporphyrin, a hypoproliferative marrow, transferrin saturation in the range of 15 to 20%, and a normal or increased serum ferritin. The serum ferritin

values are often the most distinguishing feature between true iron deficiency anemia and the iron-deficient erythropoiesis associated with inflammation. Typically, serum ferritin values increase three-fold over basal levels in the face of inflammation. All of these changes are due to the effects of inflammatory cytokines at several levels of erythropoiesis (Fig. 105-4). IL-1 directly decreases erythropoietin production in response to anemia. IL-1, acting through accessory cell release of IFN-g, suppresses the response of the erythroid marrow to erythropoietin -- an effect that can be overcome by increased erythropoietin administration in vitro and in vivo. In addition, tumor necrosis factor (TNF), acting through the release of IFN-b by marrow stromal cells, also suppresses the response to erythropoietin; several of these same cytokines, acting in concert, block the release of iron fromREstorage sites. The overall result is a chronic hypoproliferative anemia with classic changes in iron metabolism. The anemia is further compounded by a mild to moderate shortening in red cell survival.

With chronic inflammation/infection, the primary disease will determine the severity and characteristics of the anemia. For instance, many patients with cancer also have anemia that is typically normocytic and normochromic. In contrast, patients with long-standing active rheumatoid arthritis or chronic infections such as tuberculosis will have a microcytic, hypochromic anemia. In both cases, the bone marrow is hypoproliferative, but the differences in red cell indices reflect differences in the availability of iron for hemoglobin synthesis. Occasionally, conditions associated with chronic inflammation are also associated with chronic blood loss. Under these circumstances, a bone marrow aspirate stained for iron may be necessary to rule out absolute iron deficiency. However, the administration of iron in this case will correct the iron deficiency component of the anemia and leave the inflammatory component unaffected.

The anemia associated with acute infection or inflammation is typically mild, but becomes more pronounced over time. Acute infection can produce a fall in hemoglobin levels of 2 to 3 g/dL within 1 or 2 days; this is largely related to the hemolysis of red cells near the end of their natural life span. The fever and cytokines released exert a selective pressure against cells with more limited capacity to maintain the red cell membrane. In most individuals the mild anemia is reasonably well tolerated, and symptoms, if present, are associated with the underlying disease. Occasionally, in patients with preexisting cardiac disease, moderate anemia (hemoglobin 10-11 g/dL) may be associated with angina, exercise intolerance, and shortness of breath. The red cell indices vary from normocytic, normochromic to microcytic, hypochromic. The serum iron values tend to correlate with the red cell indices. The erythropoietic profile that distinguishes the anemia of inflammation from the other causes of hypoproliferative anemias is shown inTable 105-6.

ANEMIA OF RENAL DISEASE

Chronic renal failure is usually associated with a moderate to severe hypoproliferative anemia; the level of the anemia correlates with the severity of the renal failure. Red cells are typically normocytic and normochromic. Reticulocytes are decreased. The anemia is due to a failure to produce adequate amounts of erythropoietin and a reduction in red cell survival. In certain forms of acute renal failure, the correlation between the anemia and renal function is weaker. Patients with the hemolytic-uremic syndrome increase erythropoiesis in response to the hemolysis, despite renal failure requiring dialysis.

Polycystic renal disease also shows a smaller degree of erythropoietin deficiency for a given level of renal failure. By contrast, patients with diabetes have more severe erythropoietin deficiency for a given level of renal failure.

Assessment of iron status provides information to distinguish the anemia of renal disease from the other forms of hypoproliferative anemia (Table 105-6) and to guide management. Patients with the anemia of renal disease usually present with normal serum iron, TIBC, and ferritin levels. However, those maintained on chronic hemodialysis may develop iron deficiency from blood loss through the dialysis procedure. Iron must be replenished in these patients to ensure an adequate response to erythropoietin therapy (see below).

ANEMIA IN HYPOMETABOLIC STATES

Patients who are starving, particularly for protein, and those with a variety of endocrine disorders that produce lower metabolic rates may develop a mild to moderate hypoproliferative anemia. The release of erythropoietin from the kidney is sensitive to the need for O_2, not just O_2 levels. Thus, erythropoietin production is triggered at lower levels of O_2 tension in disease states (such as hypothyroidism and starvation) where metabolic activity and thus O_2 demand is decreased.

Endocrine Deficiency States The difference in the levels of hemoglobin between men and women is related to the effects of androgen and estrogen on erythropoiesis. Testosterone and anabolic steroids augment erythropoiesis; castration and estrogen administration to males decrease erythropoiesis. Patients who are hypothyroid or have deficits in pituitary hormones also may develop a mild anemia. Pathogenesis may be complicated by other nutritional deficiencies as iron and folic acid absorption can be affected by these disorders. Usually, correction of the hormone deficiency reverses the anemia.

Anemia may be more severe in Addison's disease, depending on the level of thyroid and androgen hormone dysfunction; however, anemia may be masked by decreases in plasma volume. Once such patients are given cortisol and volume replacement, the hemoglobin level may fall rapidly. Mild anemia complicating hyperparathyroidism may be due to decreased erythropoietin production as a consequence of the renal effects of hypercalcemia or to impaired proliferation of erythroid progenitors.

Protein Starvation Decreased dietary intake of protein may lead to mild to moderate hypoproliferative anemia; this form of anemia may be prevalent in the elderly. The anemia can be more severe in patients with a greater degree of starvation. In marasmus, where patients are both protein- and calorie-deficient, the release of erythropoietin is impaired in proportion to the reduction in metabolic rate; however, the degree of anemia may be masked by volume depletion and becomes apparent after refeeding. Deficiencies in other nutrients (iron, folate) may also complicate the clinical picture but may not be apparent at diagnosis. Changes in the erythrocyte indices on refeeding should prompt evaluation of iron, folate, and B_{12} status.

Anemia in Liver Disease A mild hypoproliferative anemia may develop in patients with chronic liver disease from nearly any cause. The peripheral blood smear may show burr

cells and stomatocytes from the accumulation of excess cholesterol in the membrane from a deficiency of lecithin cholesterol acyltransferase. Red cell survival is shortened, and the production of erythropoietin is inadequate to compensate. In alcoholic liver disease, nutritional deficiencies can add complexity to the management. Folate deficiency from inadequate intake and iron deficiency from blood loss and inadequate intake can alter the red cell indices.

TREATMENT

Many patients with hypoproliferative anemias experience recovery of normal hemoglobin levels when the underlying disease is appropriately treated. For those in whom such reversals are not possible -- such as patients with end-stage renal failure, cancer, and chronic inflammatory diseases -- symptomatic anemia requires treatment. The two major forms of treatment are transfusions and erythropoietin.

Transfusions Thresholds for transfusion should be altered based on the patient's symptoms. In general, patients without serious underlying cardiovascular or pulmonary disease can tolerate hemoglobin levels above 8 g/dL and do not require intervention until the hemoglobin falls below that level. Patients with more physiologic compromise may need to have their hemoglobin levels kept above 11 g/dL. A typical unit of packed red cells increases the hemoglobin level by 1 g/dL. Transfusions are associated with certain infectious risks (Chap. 114) and chronic transfusions can produce iron overload.

Erythropoietin Erythropoietin is particularly useful in anemias in which endogenous erythropoietin levels are inappropriately low, such as the hypoproliferative anemias. Iron status must be evaluated and iron repleted to obtain optimal effects from erythropoietin. In patients with chronic renal failure, the usual dose of erythropoietin is 50 to 150 U/kg three times a week subcutaneously. The dose needed to correct the anemia in patients with cancer is higher, up to 300 U/kg three times a week. Hemoglobin levels of 10 to 12 g/dL are usually reached within 4 to 6 weeks if iron levels are adequate. Once a target hemoglobin level is reached, the erythropoietin dose can be decreased to 75 U/kg three times a week. A fall in hemoglobin level occurring in the face of erythropoietin therapy usually signifies the development of an infection or iron depletion. Aluminum toxicity and hyperparathyroidism can also compromise the erythropoietin response. When an infection intervenes, it is best to interrupt the erythropoietin therapy and rely on transfusion to correct the anemia until the infection is adequately treated.

ACKNOWLEDGEMENT
Dr. Robert S. Hillman was the author of this chapter in the 14th edition, and material from his chapter has been retained.

(Bibliography omitted in Palm version)

Back to Table of Contents

106. HEMOGLOBINOPATHIES - *Edward J. Benz, Jr.*

Hemoglobin is critical for normal oxygen delivery to tissues; it is also present in erythrocytes in such high concentrations that it can alter red cell shape, deformability, and viscosity. Hemoglobinopathies are disorders affecting the structure, function, or production of hemoglobin. These disorders are usually inherited and range in severity from asymptomatic laboratory abnormalities to death in utero. Different forms may present as hemolytic anemia, erythrocytosis, cyanosis, or vasoocclusive stigmata.

PROPERTIES OF THE HUMAN HEMOGLOBINS

HEMOGLOBIN STRUCTURE

Different hemoglobins are produced during embryonic, fetal, and adult life (Fig. 106-1). Each consists of a tetramer of globin polypeptide chains: a pair of a-like chains 141 amino acids long and a pair of b-like chains 146 amino acids long. The major adult hemoglobin, HbA, has the structurea_2b_2. HbF (a_2g_2) predominates during most of gestation, and HbA_2(a_2d_2) is a minor adult hemoglobin.

Each globin chain enfolds a single heme moiety, consisting of a protoporphyrin IX ring complexed with a single iron atom in the ferrous state (Fe^{2+}), positioned in a manner optimal for reversible binding of oxygen. Each heme moiety can bind a single oxygen molecule; every molecule of hemoglobin can thus transport up to four oxygen molecules.

The amino acid sequences of the various globins are highly homologous to one another. Each has a highly helical *secondary structure*. Their globular *tertiary structures* cause the exterior surfaces to be rich in polar (hydrophilic) amino acids that enhance solubility and the interior to be lined with nonpolar groups, forming a hydrophobic "pocket" into which heme is inserted. The tetrameric *quaternary structure* of HbA contains twoab dimers. Numerous tight interactions (i.e.,$a_1b_1$contacts) hold the a and bchains together. The complete tetramer is held together by interfaces (i.e.,$a_1b_2$contacts) between the a-like chain of one dimer and the non-a chain of the other dimer.

The hemoglobin tetramer is highly soluble, but individual globin chains are insoluble. Unpaired globin precipitates, forming inclusions (Heinz bodies) that damage the cell. Normal globin chain synthesis is balanced so that each newly synthesizeda or non-a globin chain will have an available partner with which to pair to form hemoglobin.

Solubility and reversible oxygen binding are the key properties deranged in hemoglobinopathies. Both depend most on the hydrophilic surface amino acids, the hydrophobic amino acids lining the heme pocket, a key histidine in the F helix, and the amino acids forming the$a_1b_1$and$a_1b_2$contact points. Mutations in these strategic regions tend to be the ones that alter clinical behavior.

FUNCTION OF HEMOGLOBIN

To support oxygen transport, hemoglobin must bind O_2efficiently at the partial pressure of oxygen (PO_2) of the alveolus, retain it, and release it to tissues at the PO_2 of tissue

capillary beds. Oxygen acquisition and delivery over a relatively narrow range of oxygen tensions depend on a property inherent in the tetrameric arrangement of heme and globin subunits within the hemoglobin molecule called *cooperativity* or *heme-heme interaction*.

At low oxygen tensions, the hemoglobin tetramer is fully deoxygenated (Fig. 106-2). Oxygen binding begins slowly as O_2tension rises. However, as soon as some oxygen has been bound by the tetramer, an abrupt increase occurs in the slope of the curve. Thus, hemoglobin molecules that have bound some oxygen develop a higher oxygen affinity, greatly accelerating their ability to combine with more oxygen. This S-shaped oxygen equilibrium curve, along which substantial amounts of oxygen loading *and unloading* can occur over a narrow range of oxygen tensions, is physiologically more useful than the high-affinity hyperbolic curve of individual monomers.

Oxygen affinity is modulated by several factors. The Bohr effect arises from the stabilizing action of protons on deoxyhemoglobin, which binds protons more readily than oxyhemoglobin because it is a weaker acid. Thus, hemoglobin has a lower oxygen affinity at low pH, facilitating delivery to tissues (Fig. 106-2). The major small molecule that alters oxygen affinity in humans is 2,3-bisphosphoglycerate (2,3-BPG, formerly 2,3-DPG), which lowers oxygen affinity when bound to hemoglobin. HbA has a reasonably high affinity for 2,3-BPG. HbF does not bind 2,3-BPG, so it tends to have a higher oxygen affinity in vivo. Hemoglobin may also bind nitric oxide reversibly, thereby contributing to vascular tone.

To understand hemoglobinopathies, it is sufficient to understand that proper oxygen transport depends on the tetrameric structure of the proteins, the proper arrangement of the charged amino acids, and interaction with low-molecular-weight substances such as protons or2,3-BPG.

DEVELOPMENTAL BIOLOGY

Red cells first appearing at about 6 weeks after conception contain the embryonic hemoglobins Hb Portland (z_2g_2), Hb Gower I (z_2e_2), and Hb Gower II (a_2e_2). At 10 to 11 weeks, fetal hemoglobin (HbF;a_2g_2) becomes predominant. The switch to nearly exclusive synthesis of adult hemoglobin (HbA;a_2b_2) occurs at about 38 weeks (Fig. 106-1). Fetuses and newborns therefore requirea-globin but not b-globin for normal gestation. Small amounts of HbF are produced during postnatal life. A few red cell clones called *F cells* are progeny of a small pool of immature committed erythroid precursors (BFU-e) that retain the ability to produce HbF. Profound erythroid stress, such as that seen in severe hemolytic anemias, after bone marrow transplant, or during chemotherapy, cause more of the "F potent" BFU-e to be recruited. HbF levels thus tend to rise in some patients with sickle cell anemia or thalassemia. This phenomenon is also important because it probably explains the ability of hydroxyurea to increase levels of HbF in adults. Fetal globin genes can also be partially activated after birth by agents such as butyrate, which inhibit histone deacetylase and modify the structure of chromatin.

GENETICS AND BIOSYNTHESIS OF HUMAN HEMOGLOBIN

The human hemoglobins are encoded in two tightly linked gene clusters; thea-like globin genes are clustered on chromosome 16, and the b-like genes on chromosome 11 (Fig. 106-1). The a-like cluster consists of two a-globin genes and a single copy of the z gene. The non-a gene cluster consists of a singlee gene, the Gg and Ag fetal globin genes, and the adultd and bgenes.

Important regulatory sequences flank each gene. Immediately upstream are typical promoter elements needed for the assembly of the transcription initiation complex. Sequences in the 5¢ flanking region of theg and the b genes appear to be crucial for the correct developmental regulation of these genes, while elements that function like classic enhancers and silencers are in the 3¢ flanking regions. The locus control region (LCR) elements located far upstream appear to control the overall level of expression of each cluster. These elements achieve their regulatory effects by interacting with *trans*-acting transcription factors. Some of these factors are ubiquitous (e.g., Sp1 and YY1), while others are more or less limited to erythroid cells (e.g., GATA-1, NFE-2, and EKLF). The latter also appear to modulate genes specifically expressed during erythropoiesis, such as the genes that encode the enzymes of the heme biosynthetic pathway. This is relevant since normal red blood cell (RBC) differentiation also requires the coordinated expression of the globin genes with the genes responsible for heme and iron metabolism.

CLASSIFICATION OF HEMOGLOBINOPATHIES

There are five major classes of hemoglobinopathies (Table 106-1). *Structural hemoglobinopathies* occur when mutations alter the amino acid sequence of a globin chain, altering the physiologic properties of the variant hemoglobins and producing the characteristic clinical abnormalities. The variant hemoglobins relevant to this chapter polymerize abnormally, as in sickle cell anemia, or exhibit altered solubility or oxygen-binding affinity. *Thalassemia syndromes* arise from mutations that impair production or translation of globin mRNA, leading to deficient globin chain biosynthesis. Clinical abnormalities are attributable to the inadequate supply of hemoglobin and the imbalances in the production of individual globin chains, leading to premature destruction of erythroblasts and red cells. *Thalassemic hemoglobin variants* combine features of thalassemia (e.g., abnormal globin biosynthesis) and of structural hemoglobinopathies (e.g., an abnormal amino acid sequence). Hereditary persistence of fetal hemoglobin (HPFH) is characterized by synthesis of high levels of fetal hemoglobin in adult life. *Acquired hemoglobinopathies* include modifications of the hemoglobin molecule by toxins (e.g., acquired methemoglobinemia) and abnormal hemoglobin synthesis (e.g., high levels of HbF production in preleukemia and a-thalassemia in myeloproliferative disorders).

EPIDEMIOLOGY

Hemoglobinopathies are especially common in areas where malaria is endemic. This clustering of hemoglobinopathies is assumed to reflect a selective survival advantage for the abnormal red cells, which presumably provide a less hospitable environment during the obligate intraerythrocytic stages of the parasitic life cycle. Very young children with a-thalassemia are *more* susceptible to infection with the nonlethal *Plasmodium vivax*. Thalassemia might then favor a natural "vaccination" against

infection with the more lethal *P. falciparum*.

Thalassemias are the most common genetic disorders in the world, affecting nearly 200 million people worldwide. About 15% of American blacks are silent carriers for a thalassemia;a-thalassemia trait (minor) occurs in 3% of American blacks and in 1 to 15% of persons of Mediterranean origin.b Thalassemia has a 10 to 15% incidence in individuals from the Mediterranean and Southeast Asia and 0.8% in American blacks. The number of severe cases of thalassemia in the United States is about 1000. Sickle cell disease is the most common structural hemoglobinopathy occurring in heterozygous form in about 8% of American blacks and in homozygous form in 1 in 400. Between 2 and 3% of American blacks carry a hemoglobin C allele.

INHERITANCE AND ONTOGENY

Hemoglobinopathies are autosomal "codominant" traits -- compound heterozygotes that inherit a different abnormal mutant allele from each parent exhibit composite features of each. For example, patients inheriting sickleb thalassemia exhibit features ofb thalassemia and sickle cell anemia. Thea-chain is present in HbA, HBA_2, and HbF; a-chain mutations thus cause abnormalities in all three. The a-globin hemoglobinopathies are symptomatic in utero and after birth because normal function of thea-globin gene is required throughout gestation and adult life. In contrast, infants with b-globin hemoglobinopathies tend to be asymptomatic until 3 to 9 months of age, when HbA has largely replaced HbF.

DETECTION AND CHARACTERIZATION OF HEMOGLOBINOPATHIES -- GENERAL METHODS

Electrophoretic techniques are used for routine hemoglobin analysis. Electrophoresis at pH-8.6 on cellulose acetate membranes is simple, inexpensive, and reliable for initial screening. Hemoglobins S, G, and D have the same mobility at pH-8.6. Agar gel electrophoresis at pH-6.1 in citrate buffer is often used as a complementary method because it detects different variants (S migration differs from G and D). Comparison of results obtained in each system usually allows unambiguous diagnosis, but some important variants are electrophoretically silent. These mutant hemoglobins can usually be characterized by more specialized techniques such as isoelectric focusing and/or high-pressure liquid chromatography (HPLC).

Quantitation of the hemoglobin profile is often desirable. HbA_2 is frequently elevated in b-thalassemia trait and depressed in iron deficiency. HbF is elevated inHPFH, some b thalassemia syndromes, and occasional periods of erythroid stress or marrow dysplasia. For characterization of sickle cell trait, sickle thalassemia syndromes, or hemoglobin SC disease, and for monitoring the progress of exchange transfusion therapy to lower the percentage of circulating HbS, quantitation of individual hemoglobins is also required. In most laboratories, quantitation is performed only if the test is specifically ordered.

Because some variants can comigrate with HbA or HbS (sickle hemoglobin), electrophoretic assessment should always be regarded as incomplete unless functional assays for hemoglobin sickling, solubility, or oxygen affinity are also performed, as dictated by the clinical presentation. The best sickling assays involve measurement of

the degree to which the hemoglobin becomes insoluble, or gelated, as it is deoxygenated (i.e., sickle solubility test). Unstable hemoglobins are detected by their precipitation in isopropanol or after heating to 50°C. High-O_2affinity and low-O_2affinity variants are detected by quantitating the partial pressure of oxygen at which the hemoglobin sample becomes 50% saturated with oxygen (P_{50}test). Direct tests for the percentages of carboxyhemoglobin and methemoglobin, employing spectrophotometric techniques, can readily be obtained from most clinical laboratories on an urgent basis.

Complete characterization, including amino acid sequencing or gene cloning and sequencing, is available from several investigational laboratories around the world. The advent of the polymerase chain reaction (PCR), allele-specific oligonucleotide hybridization, and automated DNA sequencing has made it possible to identify globin gene mutations in a few days.

Diagnosis is best established by recognition of a characteristic history, physical findings, peripheral blood smear morphology, and abnormalities of the complete blood cell count (e.g., profound microcytosis with minimal anemia in thalassemia trait). Laboratory evaluation identifies the specific hemoglobinopathy suspected clinically.

STRUCTURALLY ABNORMAL HEMOGLOBINS

SICKLE CELL SYNDROMES

The sickle cell syndromes are caused by a mutation in the b-globin gene that changes the sixth amino acid from glutamic acid to valine. HbS ($a_2b_{26Glu®Va1}$) polymerizes reversibly when deoxygenated to form a gelatinous network of fibrous polymers that stiffen the erythrocyte membrane, increase viscosity, and cause dehydration due to potassium leakage and calcium influx (Fig. 106-3). These changes also produce the characteristic sickle shape. Sickled cells lose the pliability needed to traverse small capillaries. They possess altered "sticky" membranes (especially reticulocytes) that are abnormally adherent to the endothelium of small venules. These abnormalities provoke unpredictable episodes of microvascular vasoocclusion and premature red cell destruction (hemolytic anemia). Hemolysis occurs because the abnormal erythrocytes are destroyed by the spleen. The rigid adherent cells also clog small capillaries and venules, causing tissue ischemia, acute pain, and gradual end-organ damage. This venoocclusive component usually dominates the clinical course. Prominent manifestations include episodes of ischemic pain (i.e., painful crises) and ischemic malfunction or frank infarction in the spleen, central nervous system, bones, liver, kidneys, and lungs.

The prototype disease, sickle cell anemia, is the homozygous state for HbS (Table 106-2). Several sickle syndromes occur as the result of inheritance of HbS from one parent and another hemoglobinopathy, such as b thalassemia or HbC ($a_2b_{26Glu®Lys}$) from the other parent.

Clinical Manifestations

Sickle Cell Anemia Most patients with sickling syndromes suffer from hemolytic anemia, with hematocrits of 15 to 30%, and significant reticulocytosis. Anemia was once thought

to exert protective effects against vasoocclusion by reducing blood viscosity. Natural history and drug therapy trials suggest that an *increase* in the hematocrit with feedback inhibition of reticulocytosis might be beneficial, even at the expense of increased blood viscosity. The role of adhesive reticulocytes in vasoocclusion might account for these paradoxical effects.

Granulocytosis is common. The white cell count can fluctuate substantially and unpredictably during and between painful crises, infectious episodes, and other intercurrent illnesses.

Vasoocclusion causes protean manifestations; intermittent episodes in connective and musculoskeletal structures produce painful ischemia manifested by acute pain and tenderness, fever, tachycardia, and anxiety. These recurrent episodes, called *painful crises*, are the most common clinical manifestation. Their frequency and severity vary greatly. Pain can develop almost anywhere in the body and may last from a few hours to 2 weeks. Repeated crises requiring hospitalization (more than three per year) correlate with reduced survival in adult life, suggesting that these episodes are associated with accumulation of chronic end-organ damage. Provocative factors include infection, fever, excessive exercise, anxiety, abrupt changes in temperature, hypoxia, or hypertonic dyes.

Repeated microinfarction can destroy tissues having microvascular beds that promote sickling. Thus, the spleen is frequently infarcted within the first 18 to 36 months of life, causing susceptibility to infection, particularly from pneumococci. Acute venous obstruction of the spleen (*splenic sequestration crisis*), a rare occurrence in early childhood, may require emergency transfusion and/or splenectomy to prevent trapping of the entire arterial output in the obstructed spleen. Occlusion of retinal vessels can produce hemorrhage, neovascularization, and eventual detachments. Renal papillary necrosis invariably produces isosthenuria. More widespread renal necrosis leads to renal failure in adults, a common late cause of death. Bone and joint ischemia can lead to aseptic necrosis (especially of the femoral or humeral heads), chronic arthropathy, and unusual susceptibility to osteomyelitis, which may be caused by organisms such as *Salmonella*, rarely encountered in other settings. The *hand-foot syndrome* is caused by painful infarcts of the digits and dactylitis. Stroke is especially common in children, a small subset of whom tend to suffer repeated episodes; stroke is less common in adults and is often hemorrhagic. A particularly painful complication in males is priapism, due to infarction of the penile venous outflow tracts; permanent impotence is a frequent consequence. Chronic lower leg ulcers probably arise from ischemia and superinfection in the distal circulation.

Acute chest syndrome is a distinctive manifestation characterized by chest pain, tachypnea, fever, cough, and arterial oxygen desaturation. It can mimic pneumonia, pulmonary emboli, bone marrow infarction and embolism, myocardial ischemia, or in situ lung infarction. Acute chest syndrome is thought to reflect in situ sickling within the lung, producing pain and temporary pulmonary dysfunction. Acute chest syndrome may be difficult or impossible to distinguish from other entities. Pulmonary infarction and pneumonia are the most frequent underlying or concomitant conditions in patients with this syndrome. Repeated episodes of acute chest pain correlate with reduced survival. Acutely, reduction in arterial oxygen saturation is especially ominous because it

promotes sickling on a massive scale. Repeated acute or subacute pulmonary crises lead to pulmonary hypertension and cor pulmonale, an increasingly common cause of death as patients survive further into adult life.

Sickle cell syndromes are remarkable for their clinical heterogeneity. Some patients remain virtually asymptomatic into or even through adult life, while others suffer repeated crises requiring hospitalization from early childhood. At least five haplotypes of sickle cell disease are recognized based upon their origin: Senegal, Cameroon, Benin, Central African Republic, and India. Among these, patients of the Central African Republic have the worst disease and those of Senegal the least severe. Patients with sickle thalassemia and sickle-HbE tend to have similar, slightly milder, symptoms, perhaps because of the ameliorating effects of production of other hemoglobins within the red cell. Hemoglobin SC disease, one of the more common variants of sickle cell anemia, is frequently marked by lesser degrees of hemolytic anemia and a greater propensity for the development of retinopathy and aseptic necrosis of bones. In most respects, however, the clinical manifestations resemble sickle cell anemia. Some rare hemoglobin variants actually aggravate the sickling phenomenon.

Sickle Cell Trait Sickle cell trait is usually asymptomatic. Anemia and painful crises are exceedingly rare. An uncommon, but highly distinctive, symptom is painless hematuria, often occurring in adolescent males, probably due to papillary necrosis. Sloughing of papillae with ureteral obstruction has been reported, as have isolated cases of massive sickling or sudden death due to exposure to high altitudes or extraordinary extremes of exercise and dehydration.

Diagnosis Sickle cell syndromes are readily suspected on the basis of characteristic hemolytic anemia, red cell morphology (Plate V-39), and intermittent episodes of ischemic pain. Diagnosis is confirmed by hemoglobin electrophoresis and sickling tests. Thorough characterization of the exact hemoglobin profile of the patient is important, because sickle thalassemia and hemoglobin SC disease are correlated with alterations in prognosis or clinical features. The diagnosis is usually established in childhood, but occasional patients, often with compound heterozygous states, do not develop symptoms until the onset of puberty, pregnancy, or early adult life. Genotyping of family members and potential parental partners is critical for genetic counseling. Details of the childhood history help to establish prognosis and eligibility for aggressive or experimental therapies. Factors associated with increased morbidity and mortality are more than three crises requiring hospitalization per year, chronic neutrophilia, a history of splenic sequestration or hand-foot syndrome, and second episodes of acute chest syndrome. Patients with a history of cerebrovascular accidents are at higher risk for repeated episodes and require especially close monitoring.

TREATMENT

Patients with sickle cell syndromes require ongoing continuity of care. Familiarity with the pattern of symptoms provides the best safeguard against excessive use of the emergency room, hospitalization, and habituation to addictive narcotics. Additional preventive measures include regular slit-lamp examinations to monitor development of retinopathy; antibiotic prophylaxis appropriate for splenectomized patients during dental or other invasive procedures; vaccination against pneumococci and *Haemophilus*

influenzae; and vigorous oral hydration before or during periods of extreme exercise, exposure to heat or cold, emotional stress, or infection.

The management of acute painful crisis includes vigorous hydration, thorough evaluation for underlying causes (such as infection), and aggressive narcotic analgesia administered by a standing order and/or PCA pump. Morphine (0.1 to 0.15 mg/kg every 3 to 4 h) or meperidine (0.75 to 1.5 mg/kg every 2 to 4 h) should control severe pain. Bone pain may respond as well to ketorolac (30 to 60 mg initial dose, then 15 to 30 mg every 6 to 8 h). Many crises can be managed at home with oral hydration and oral analgesia. Use of the emergency room should be reserved for especially severe symptoms or circumstances in which other processes (e.g., infection) are strongly suspected. Nasal oxygen should be employed as appropriate to protect arterial saturation. Most crises resolve in 1 to 7 days. Use of blood transfusion should be reserved for extreme cases; transfusion does not shorten the crisis.

No tests are definitive to diagnose acute painful crisis. Critical to good management is an approach that recognizes that most patients reporting crisis symptoms do indeed have crisis or another significant medical problem. Diligent diagnostic evaluation for underlying causes is imperative, even though these are found infrequently. In adults, the possibility of aseptic necrosis or sickle arthropathy must be considered, especially if pain and immobility become repeated or chronic at a single site. Nonsteroidal anti-inflammatory agents are often effective for sickle cell arthropathy.

Acute chest syndrome is a medical emergency that may require management in an intensive care unit. Hydration should be monitored carefully to avoid the development of pulmonary edema, and oxygen therapy should be especially vigorous for protection of arterial saturation. Diagnostic evaluation for pneumonia and pulmonary embolism should be thorough, since these may occur with atypical symptoms. Critical interventions are transfusion to maintain a hematocrit >30 and emergency exchange transfusion if arterial saturation drops below 90%.

As patients with sickle cell syndromes increasingly survive into their fifth and sixth decades (median age at death is 42 years for men, 48 years for women), end-stage renal failure and pulmonary hypertension are becoming increasingly prominent causes of end-stage morbidity; anecdotal evidence suggests that a sickle cell cardiomyopathy and/or premature coronary artery disease may compromise cardiac function in later years. Sickle cell patients have received kidney transplants, but they often experience an increase in the frequency and severity of crises, possibly due to increased infection as a consequence of immunosuppression.

The most significant advance in the therapy of sickle cell anemia has been the introduction of hydroxyurea as a mainstay of therapy for patients with severe symptoms. Hydroxyurea (10 to 30 mg/kg/per day) increases fetal hemoglobin and may also exert beneficial affects on red cell hydration, vascular wall adherence, and suppression of the granulocyte and reticulocyte counts; indeed, dosage is titrated to maintain a white cell count between 5,000 and 8,000. White cells and reticulocytes may play a major role in the pathogenesis of sickle cell crisis, and their suppression may be an important benefit of hydroxyurea therapy.

Hydroxyurea should be considered in patients experiencing repeated episodes of acute chest syndrome or more than three crises per year requiring hospitalization. The utility of this agent for reducing the incidence of other complications (e.g., priapism, retinopathy) is under evaluation, as are the long-term side effects. Therefore, when possible, treatment should be instituted as part of a clinical trial. Most patients respond within a few months with elevations of fetal hemoglobin.

Bone marrow transplantation can provide definitive cures but is known to be effective and safe only in children. Prognostic features justifying bone marrow transplant are the presence of repeated crises early in life, a high neutrophil count, or the development of hand-foot syndrome. Children at risk for stroke can be identified through the use of Doppler ultrasound techniques. Prophylactic exchange transfusion appears to reduce the risk of stroke substantially in this population. Children who do suffer a cerebrovascular accident should be maintained for at least 3 to 5 years on a program of vigorous exchange transfusion, since the risk of second strokes is extremely high in this population.

Gene therapy for sickle cell anemia is under investigation, but no safe therapy is currently available. Agents blocking red cell hydration or vascular adhesion, such as clotrimazole, may have value as an adjunct to hydroxyurea therapy; trials are ongoing.

UNSTABLE HEMOGLOBINS

Amino acid substitutions that reduce solubility or increase susceptibility to oxidation produce "unstable" hemoglobins that precipitate, forming inclusion bodies injurious to the red cell membrane. Representative mutations are those that interfere with contact points between the a and b subunits [e.g., Hb Philly ($b_{35Tyr®Phe}$)], alter the helical segments [e.g., Hb Genova ($b_{28Leu®Pro}$)], or disrupt interactions of the hydrophobic pockets of the globin subunits with heme [e.g., Hb Koln ($b_{98Val®Met}$)] (Table 106-3). The inclusions, called *Heinz bodies*, are clinically detectable by staining with supravital dyes such as crystal violet (Heinz body test). Removal of these inclusions by the spleen generates pitted, rigid cells that have shortened life spans, producing hemolytic anemia of variable severity, sometimes requiring chronic transfusion support. Splenectomy may be needed to correct the anemia. Leg ulcers and premature gallbladder disease due to bilirubin turnover are frequent stigmata.

Unstable hemoglobins occur sporadically, often by spontaneous new mutations. Heterozygotes are often symptomatic because a significant Heinz body burden can develop even when the unstable variant accounts for a portion of the total hemoglobin. Symptomatic unstable hemoglobins tend to be b-globin variants, because sporadic mutations affecting only one of the foura globins would generate only 20 to 30% abnormal hemoglobin.

HEMOGLOBINS WITH ALTERED OXYGEN AFFINITY

High-affinity hemoglobins [e.g., Hb Yakima ($b_{99Asp®His}$)] bind oxygen more readily but deliver less O_2 to tissues at normal capillary P_{O2}levels (Fig. 106-2). Mild tissue hypoxia ensues, stimulatingRBCproduction and erythrocytosis (Table 106-3). In extreme cases, the hematocrit can rise to 60 to 65%, increasing blood viscosity and producing typical

symptoms (headache, somnolence, or dizziness). Phlebotomy may be required. Typical mutations alter interactions within the heme pocket or disrupt the Bohr effect or salt-bond site. Mutations that impair the interaction of HbA with2,3-BPGcan increase O_2affinity, because 2,3-BPG binding lowers O_2affinity.

Low-affinity hemoglobins [e.g., Hb Kansas ($b_{102Asn®Thr}$)] bind sufficient oxygen in the lungs, despite their lower oxygen affinity, to achieve nearly full saturation. At capillary oxygen tensions, they lose sufficient amounts of oxygen to maintain homeostasis at a low hematocrit (Fig. 106-2) (pseudoanemia). Capillary hemoglobin desaturation can also be sufficient to produce clinically apparent cyanosis. Despite these findings, patients usually require no specific treatment.

METHEMOGLOBINEMIAS

Methemoglobin is generated by oxidation of the heme iron moieties to the ferric state, causing a characteristic bluish-brown, muddy color resembling cyanosis. Methemoglobin has such high oxygen affinity that virtually no oxygen is delivered to tissues. Levels >50 to 60% are often fatal.

Congenital methemoglobinemia arises from globin mutations that stabilize iron in the ferric state [e.g., HbM Iwata ($a_{87His®Tyr}$),Table 106-3] or from mutations that impair the enzymes that reduce methemoglobin to hemoglobin (e.g., methemoglobin reductase, NADP diaphorase). Acquired methemoglobinemia is caused by toxins that oxidize heme iron, notably nitrate and nitrite-containing compounds.

DIAGNOSIS AND MANAGEMENT OF PATIENTS WITH UNSTABLE HEMOGLOBINS, HIGH-AFFINITY HEMOGLOBINS, AND METHEMOGLOBINEMIA

Unstable hemoglobin variants should be suspected in patients with nonimmune hemolytic anemia, jaundice, splenomegaly, or premature biliary tract disease. Severe hemolysis usually presents during infancy as neonatal jaundice or anemia. Milder cases may present in adult life with anemia or only as unexplained reticulocytosis, hepatosplenomegaly, premature biliary tract disease, or leg ulcers. Because spontaneous mutation is common, family history of anemia may be absent. The peripheral blood smear often shows anisocytosis, abundant cells with punctate inclusions, and irregular shapes (i.e., poikilocytosis).

The two best tests for diagnosing unstable hemoglobins are the Heinz body preparation and the isopropanol or heat stability test. Many unstable Hb variants are electrophoretically silent. A normal electrophoresis does not rule out the diagnosis.

Severely affected patients may require transfusion support for the first 3 years of life, because splenectomy before age 3 is associated with a significantly greater immune deficit. Splenectomy is usually effective thereafter, but occasional patients may require lifelong transfusion support. Even after splenectomy, patients can develop cholelithiasis and leg ulcers. Splenectomy can also be considered in patients exhibiting severe secondary complications of chronic hemolysis, even if anemia is absent. Precipitation of unstable hemoglobins is aggravated by oxidative stress, e.g., infection, antimalarial drugs.

High-O_2-affinity hemoglobin variants should be suspected in patients with erythrocytosis. The best test for confirmation is measurement of the P_{50}. A high-O_2-affinity Hb causes a significant left shift (i.e., lower numeric value of the P_{50}); confounding conditions, e.g., tobacco smoking or carbon monoxide exposure, can also lower the P_{50}.

Patients with high-affinity hemoglobin are often asymptomatic; rubor or plethora may be telltale signs. When the hematocrit reaches 55 to 60%, symptoms of high blood viscosity and sluggish flow (headache, lethargy, dizziness, etc.) may be present. These symptoms respond to judicious phlebotomy. Erythrocytosis represents an appropriate attempt to compensate for the impaired oxygen delivery by the abnormal variant. Overzealous phlebotomy may stimulate increased erythropoiesis or aggravate symptoms by thwarting this compensatory mechanism. The guiding principle of phlebotomy should be to improve oxygen delivery by reducing blood viscosity and increasing blood flow rather than restoration of a normal hematocrit. Modest iron deficiency may aid in control.

Low-affinity hemoglobins should be considered in patients with cyanosis or a low hematocrit with no other cause apparent after thorough evaluation. The P_{50} test confirms the diagnosis. Counseling and reassurance are the interventions of choice.

Methemoglobin should be suspected in patients with hypoxic symptoms who appear cyanotic but have a Pa_{O2}sufficiently high that hemoglobin should be fully saturated with oxygen. A history of nitrite or other oxidant ingestions may not always be available; some exposures may be unapparent to the patient, and others may result from suicide attempts. The characteristic muddy appearance of freshly drawn blood can be a critical clue. The diagnostic test of choice is measurement of the methemoglobin content, which is usually available on an emergency basis.

Methemoglobinemia often causes symptoms of cerebral ischemia at levels>15%; levels>60% are usually lethal. Intravenous injection of 1 mg/kg of methylene blue is effective emergency therapy. Milder cases and follow-up of severe cases can be treated orally with methylene blue (60 mg three to four times each day) or ascorbic acid (300 to 600 mg/d).

THALASSEMIA SYNDROMES

The thalassemia syndromes are inherited disorders of a- orb-globin biosynthesis. The reduced supply of globin diminishes production of hemoglobin tetramers, causing hypochromia and microcytosis. Unbalanced accumulation of a andb subunits occurs because the synthesis of the unaffected globins proceeds at normal rate. Unbalanced chain accumulation dominates the clinical phenotype. Clinical severity varies widely, depending on the degree to which the synthesis of the affected globin is impaired, altered synthesis of other globin chains, and coinheritance of other abnormal globin alleles.

b-THALASSEMIA SYNDROMES

Mutations causing thalassemia can affect any step in the pathway of globin gene

expression: transcription, processing of the mRNA precursor, translation, and posttranslational metabolism of the b-globin polypeptide chain. The most common forms arise from mutations that derange splicing of the mRNA precursor or prematurely terminate translation of the mRNA.

Hypochromia and microcytosis due to reduced amounts of hemoglobin tetramers characterize all forms ofb thalassemia. In heterozygotes (b-thalassemia trait), this is the only abnormality seen; anemia is minimal. In homozygous states, unbalanceda- and b-globin accumulation causes accumulation of highly insoluble unpaired achains, which form toxic inclusion bodies that kill developing erythroblasts in the marrow. Few of the proerythroblasts beginning erythroid maturation survive. The few surviving red cells bear a burden of inclusion bodies, detected in the spleen, shortening the red cell life span and producing severe hemolytic anemia. The resulting profound anemia stimulates erythropoietin release and compensatory erythroid hyperplasia, but the marrow response is sabotaged by ineffective erythropoiesis. Anemia persists. Erythroid hyperplasia can become exuberant and produce extramedullary erythropoietic tissue in the liver and spleen.

Massive bone marrow expansion deranges growth and development. Children develop characteristic "chipmunk" facies due to maxillary marrow hyperplasia and frontal bossing, thinning and pathologic fracture of long bones and vertebrae due to cortical invasion by erythroid elements, and profound growth retardation. Hemolytic anemia causes hepatosplenomegaly, leg ulcers, gallstones, and high-output congestive heart failure. The conscription of caloric resources to support erythropoiesis leads to inanition, susceptibility to infection, endocrine dysfunction, and, in the most severe cases, death during the first decade of life. Chronic transfusions with red cells improves oxygen delivery, suppresses the excessive ineffective erythropoiesis, and prolongs life, but the inevitable side effects, notably iron overload, usually prove fatal by age 30. Bone marrow transplantation in childhood is the only curative therapy.

Severity is highly variable. Known modulating factors are those that ameliorate the burden of unpaireda-globin inclusions. Alleles associated with milder synthetic defects and coinheritance of a-thalassemia trait reduce clinical severity by reducing accumulation of excessa globin. HbF persists to various degrees inb thalassemias. g-Globin gene chains can substitute for b chains, simultaneously generating more hemoglobin and reducing the burden ofa-globin inclusions. The terms*b-thalassemia major* and*b-thalassemia intermedia* are used to reflect the clinical heterogeneity. Patients withb-thalassemia major require intensive transfusion support to survive. Patients with b-thalassemia intermedia have a somewhat milder phenotype and can survive without transfusion. The terms *b-thalassemia minor* and*b-thalassemia trait* describe asymptomatic heterozygotes for bthalassemia.

a-THALASSEMIA SYNDROMES

The four classica thalassemias, most common in Asians, area-thalassemia-2 trait, in which one of the foura-globin loci is deleted;a-thalassemia-1 trait, with two deleted loci; HbH disease, with three loci deleted; and hydrops fetalis with Hb Bart's, with all four loci deleted (Table 106-4). Nondeletion forms ofa thalassemia also exist.

a-Thalassemia-2 trait is an asymptomatic, silent carrier state.*a-Thalassemia-1 trait* resemblesb-thalassemia minor. Offspring doubly heterozygous fora-thalassemia-2 anda-thalassemia-1 exhibit a more severe phenotype, called HbH disease. Heterozygosity for a deletion that removes both genes from the same chromosome (*cis* deletion) is common in Asians and Mediterranean individuals, as is homozygosity for a-thalassemia-2 (*trans* deletion). Both produce asymptomatic hypochromia and microcytosis.

In *HbH disease*, HbA production is only 25 to 30% of normal. Fetuses accumulate some unpaired b chains. In adults, unpairedb chains accumulate and are soluble enough to form$b_4$tetramers called *HbH*. HbH forms few inclusions in erythroblasts but does precipitate in circulating red cells. Patients with HbH disease have thalassemia intermedia characterized by moderately severe hemolytic anemia but milder ineffective erythropoiesis. Survival into midadult life without transfusions is common.

The homozygous state for thea-thalassemia-1 *cis* deletion (hydrops fetalis) causes total absence of a-globin synthesis. No physiologically useful hemoglobin is produced beyond the embryonic stage. Excessg globin forms tetramers called *Hb Bart's* (g_4), which has an extraordinarily high oxygen affinity. It delivers almost no O_2 to fetal tissues, causing tissue asphyxia, edema (hydrops fetalis), congestive heart failure, and death in utero. a-Thalassemia-2 trait is common (15 to 20%) among people of African descent. The *cis* a-thalassemia-1 deletion is almost never seen, however. Thus, a-thalassemia-2 and the *trans* form of a-thalassemia-1 are very common, but HbH disease and hydrops fetalis are almost never encountered.

DIAGNOSIS AND MANAGEMENT

The diagnosis ofb-thalassemia major is readily made during childhood on the basis of severe anemia accompanied by hepatosplenomegaly; profound microcytosis; a characteristic blood smear (Plate V-2); and elevated levels of HbF, HbA_2, or both. Many patients require chronic hypertransfusion therapy designed to maintain a hematocrit of at least 27 to 30% so that erythropoiesis is suppressed. Splenectomy is required if the annual transfusion requirement (volume ofRBCs per kilogram body weight per year) increases by>50%. Folic acid supplements may be useful. Vaccination with pneumococcal vaccine in anticipation of eventual splenectomy is advised, as is close monitoring for infection, leg ulcers, and biliary tract disease. Early endocrine evaluation is required for glucose intolerance, thyroid dysfunction, and delayed onset of puberty or secondary sexual characteristics. Many patients develop endocrine deficiencies as a result of iron overload.

Patients withb-thalassemia intermedia exhibit similar stigmata but can survive without chronic hypertransfusion. Management is particularly challenging because a number of factors can aggravate the anemia, including infection, onset of puberty, and development of splenomegaly and hypersplenism. Some patients may eventually benefit from splenectomy. The expanded erythron can cause excess absorption of dietary iron and hemosiderosis, even without transfusion.

b-Thalassemia minor (i.e., thalassemia trait) usually presents as profound microcytosis and hypochromia with target cells but only minimal or mild anemia. The mean

corpuscular volume is rarely >75 fL; the hematocrit is rarely<30 to 33%. Hemoglobin electrophoresis classically reveals an elevated HbA_2(3.5 to 7.5%), but some forms are associated with normal $HbA_2$and/or elevated HbF. Genetic counseling and patient education are essential. Patients with b-thalassemia trait should be warned that their blood picture resembles iron deficiency and can be misdiagnosed. They should eschew routine use of iron but know that iron deficiency requiring supplementation can develop, as in other persons, during pregnancy or from chronic bleeding.

Persons witha-thalassemia trait may exhibit mild hypochromia and microcytosis, usually without anemia. HbA_2 and HbF levels are normal. Affected individuals usually require only genetic counseling. HbH disease resembles b-thalassemia intermedia, with the added complication that the HbH molecule behaves like a moderately unstable hemoglobin. Patients with HbH disease should undergo splenectomy if excessive anemia or a transfusion requirement develops. Oxidative drugs should be avoided. Iron overload leading to death can occur in more severely affected patients.

PREVENTION

Antenatal diagnosis of thalassemia syndromes is now widely available. DNA diagnosis is based onPCRamplification of fetal DNA, obtained by amniocentesis or chorionic villus biopsy followed by hybridization to allele-specific oligonucleotides probes. The probes can be designed to detect simultaneously the subset of mutations that account for 95 to 99% of the a orb thalassemias that occur in a particular ethnic group.

THALASSEMIC STRUCTURAL VARIANTS

Thalassemic structural variants are characterized by both defective synthesis and abnormal structure.

HEMOGLOBIN LEPORE

Hb Lepore [$a_2(db)_2$] arises by an unequal crossover and recombination event that fuses the proximal end of the d gene with the distal end of the closely linked b gene. The resulting chromosome contains only the fused db gene. The Lepore (db) globin is synthesized poorly because the fused gene is under the control of the weak d-globin promoter. Hb Lepore alleles have a phenotype likeb-thalassemia, except for the added presence of 2 to 20% Hb Lepore. Compound heterozygotes for Hb Lepore and a classic b-thalassemia allele may also have severe thalassemia.

HEMOGLOBIN E

HbE (i.e.,$a_2b_{226Glu®Lys}$) is extremely common in Cambodia, Thailand, and Vietnam. The gene has become far more prevalent in the United States as a result of immigration of Asian persons, especially in California, where HbE is the most common variant detected. HbE is mildly unstable but not enough to affectRBC life span significantly. The high frequency of the HbE gene may be a result of the thalassemia phenotype associated with its inheritance. Heterozygotes resemble individuals with mild b-thalassemia trait. Homozygotes have somewhat more marked abnormalities but are asymptomatic. Compound heterozygotes for HbE and a b-thalassemia gene can have

b-thalassemia intermedia orb-thalassemia major, depending on the severity of the coinherited thalassemic gene.

Theb^{E}allele contains only a single base change, in codon 26, that causes the amino acid substitution. However, this mutation activates a cryptic RNA splice site generating a structurally abnormal globin mRNA that cannot be translated from about 50% of the initial pre-mRNA molecules. The remaining 40 to 50%, which are normally spliced, generate functional mRNA that is translated intob^{E}globin because the mature mRNA carries the base change that alters codon 26.

Genetic counseling of the persons at risk for HbE should focus on the interaction of HbE withb-thalassemia rather than HbE homozygosity, a condition associated with microcytosis and hypchromia that is usually asymptomatic, with hemoglobin levels rarely <10 gm/dL.

OTHER UNCOMMON HEMOGLOBINOPATHIES

HEREDITARY PERSISTENCE OF FETAL HEMOGLOBIN

HPFHis characterized by continued synthesis of high levels of HbF in adult life. No deleterious effects are apparent, even when all of the hemoglobin produced is HbF. These rare patients demonstrate convincingly that prevention or reversal of the fetal to adult hemoglobin switch would provide efficacious therapy for sickle cell anemia and bthalassemia.

ACQUIRED HEMOGLOBINOPATHIES

The two most important acquired hemoglobinopathies are carbon monoxide poisoning and methemoglobinemia (see above). Carbon monoxide has a higher affinity for hemoglobin than does oxygen; it can replace oxygen and diminish O_2delivery. Chronic elevation of carboxyhemoglobin levels to 10 or 15%, as occurs in smokers, can lead to secondary polycythemia. Carboxyhemoglobin is cherry red in color and masks the development of cyanosis usually associated with poor O_2delivery to tissues.

Abnormalities of hemoglobin biosynthesis have also been described in blood dyscrasias. In some patients with myelodysplastic, erythroleukemic, or myeloproliferative disorders, a mild form of HbH disease may also be seen. The abnormalities are not severe enough to alter the course of the underlying disease.

MANAGEMENT OF TRANSFUSIONAL HEMOSIDEROSIS

Chronic blood transfusion can lead to blood-borne infection, alloimmunization, febrile reactions, and lethal iron overload. A unit of packedRBCs contains 250 to 300 mg iron (1 mg/mL). The iron assimilated by a single transfusion of two units of packed RBCs is thus equal to a 1- to 2-year intake of iron. Iron accumulates in chronically transfused patients because no mechanisms exist for increasing iron excretion; an expanded erythron causes especially rapid development of iron overload because accelerated erythropoiesis promotes excessive absorption of dietary iron. Vitamin C should not be supplemented because it generates free radicals in iron excess states.

Patients who receive>100 units of packedRBCs usually develop hemosiderosis. The ferritin level rises, followed by early endocrine dysfunction (glucose intolerance and delayed puberty), cirrhosis, and cardiomyopathy. Liver biopsy shows both parenchymal and reticuloendothelial iron. Newer methods for assessing hepatic iron such as the superconducting quantum-interference device (SQUID) are accurate but not widely available. Cardiac toxicity is often insidious. Early development of pericarditis is followed by dysrhythmia and pump failure. The onset of heart failure is ominous, often presaging death within a year (Chap. 345).

The decision to start long-term transfusion support should be accompanied by therapy with iron-chelating agents. The only approved and available iron chelator, desferoxamine (Desferal), is expensive and poorly absorbed from the gastrointestinal tract. Its iron-binding kinetics require chronic slow infusion via a metering pump. The constant presence of the drug improves the efficiency of chelation and protects tissues from occasional releases of the most toxic fraction of iron -- low-molecular-weight iron -- which may not be sequestered by protective proteins. Oral iron-chelating agents such as deferiprone showed initial promise, but long-term trials have raised serious doubts about their efficacy and safety.

Desferoxamine is relatively nontoxic. Occasional cataracts, deafness, and local skin reactions, including urticaria, occur. Skin reactions can usually be managed with antihistamines. Negative iron balance can be achieved, even in the face of a high transfusion requirement, but this alone does not prevent long-term morbidity and mortality in chronically transfused patients. Irreversible end-organ deterioration develops at relatively modest levels of iron overload, even if symptoms do not appear for many years thereafter. To obtain a significant survival advantage, chelation must begin before 5 to 8 years of age.

EXPERIMENTAL THERAPIES

Bone marrow transplantation provides stem cells able to express normal hemoglobin; it has been used in a large number of patients with b thalassemia and a smaller number of patients with sickle cell anemia. Early in the course of disease, before end-organ damage occurs, transplantation is curative in 80 to 90% of patients. In highly experienced centers, the treatment-related mortality is<10%. Since survival into adult life is possible with conventional therapy, the decision to transplant is best made in consultation with specialized centers.

Gene therapy of thalassemia and sickle cell disease has proved to be an elusive goal. Uptake of gene vectors into the nondividing hematopoietic stem cells has been disappointingly inefficient.

Reestablishing high levels of fetal hemoglobin synthesis should ameliorate the symptoms ofb thalassemia. Cytotoxic agents such as hydroxyurea and cytarabine promote high levels of HbF synthesis, probably by stimulating proliferation of the primitive HbF-producing progenitor cell population (i.e., F cell progenitors). Unfortunately, no regimen has yet been identified that ameliorates the clinical manifestations of bthalassemia. Butyrates stimulate HbF production, but only

transiently. Pulsed or intermittent administration has been found to sustain HbF induction in the majority of patients with sickle cell disease. It is unclear whether butyrate will have similar activity in patients with bthalassemia.

APLASTIC AND HYPOPLASTIC CRISIS IN PATIENTS WITH HEMOGLOBINOPATHIES

Patients with hemolytic anemia sometimes exhibit an alarming decline in hematocrit during and immediately after acute illnesses. Bone marrow suppression occurs in almost everyone during acute inflammatory illnesses. In patients with shortened red cell life spans, suppression can cause anemia. These hypoplastic crises are usually transient and do not require transfusion.

Aplastic crisis refers to a profound cessation of erythroid activity in patients with chronic hemolytic anemia. It is associated with a rapidly falling hematocrit. Episodes are usually self-limited. Aplastic crises are caused by infection with a particular strain of parvovirus (B19A). Children infected with this virus usually develop permanent immunity. Aplastic crises do not often recur and are rarely seen in adults. Management requires close monitoring of the hematocrit and reticulocyte count. If anemia becomes symptomatic, transfusion support is indicated. Most crises resolve spontaneously within 1 to 2 weeks.

ACKNOWLEDGEMENT

Some material from Chap. 107 by Dr. Ernest Beutler in the last edition has been retained in this edition. In addition, portions of this chapter describe well-established aspects of this topic and are revised and updated from earlier chapters on this topic by the author.

(Bibliography omitted in Palm version)

Back to Table of Contents

107. MEGALOBLASTIC ANEMIAS - *Bernard M. Babior*, *H. Franklin Bunn*

The megaloblastic anemias are disorders caused by impaired DNA synthesis. Cells primarily affected are those having relatively rapid turnover, especially hematopoietic precursors and gastrointestinal epithelial cells. Cell division is sluggish, but cytoplasmic development progresses normally, so megaloblastic cells tend to be large, with an increased ratio of RNA to DNA. Megaloblastic erythroid progenitors tend to be destroyed in the marrow. Thus, marrow cellularity is often increased but production of red blood cells (RBC) is decreased, an abnormality termed *ineffective erythropoiesis* (Chap. 61).

Most megaloblastic anemias are due to a deficiency of cobalamin (vitamin B_{12}) and/or folic acid. The various clinical entities associated with megaloblastic anemia are listed in Table 107-1.

PHYSIOLOGIC AND BIOCHEMICAL CONSIDERATIONS

FOLIC ACID

Folic acid is the common name for pteroylmonoglutamic acid. It is synthesized by many different plants and bacteria. Fruits and vegetables constitute the primary dietary source of the vitamin. Some forms of dietary folic acid are labile and may be destroyed by cooking. The minimum daily requirement is normally about 50 ug, but this may be increased severalfold during periods of enhanced metabolic demand such as pregnancy.

The assimilation of adequate amounts of folic acid depends on the nature of the diet and its means of preparation. Folates in various foodstuffs are largely conjugated to a chain of glutamic acid residues. This highly polar side chain impairs the intestinal absorption of the vitamin. However, conjugases (g-glutamyl carboxypeptidases) in the lumen of the gut convert polyglutamates to mono- and diglutamates, which are readily absorbed in the proximal jejunum.

Plasma folate is primarily in the form of N_5-methyltetrahydrofolate, a monoglutamate, which is transported into cells by a carrier that is specific for the tetrahydro forms of the vitamin. Once in the cell, the N_5-methyl group is removed in a cobalamin-requiring reaction (see below), and the folate is then reconverted to the polyglutamate form. The polyglutamate form may be useful for retention of folate by the cell.

A folate-binding protein occurs in plasma, milk, and other body fluids. The function of this folate binder and its membrane-bound precursor is unknown. Neither the binder nor its precursor is related to the tetrahydrofolate carrier.

Normal individuals have about 5 to 20 mg folic acid in various body stores, half in the liver. In light of the minimum daily requirement, it is not surprising that a deficiency will occur within months if dietary intake or intestinal absorption is curtailed.

The prime function of folate compounds is to transfer 1-carbon moieties such as methyl and formyl groups to various organic compounds (Fig. 107-1). The sources of these

1-carbon moieties is usually serine, which reacts with tetrahydrofolate to produce glycine and $N_{5,10}$-methylenetetrahydrofolate. An alternative source is forminoglutamic acid, an intermediate in histidine catabolism, which gives up its formimino group to tetrahydrofolate to yield N_5-formiminotetrahydrofolate and glutamic acid. These derivatives provide entry into an interconvertible donor pool consisting of tetrahydrofolate derivatives carrying various 1-carbon moieties. The constituents of this pool can donate their 1-carbon moieties to appropriate acceptor compounds to form metabolic intermediates, which are ultimately converted to building blocks used in the synthesis of macromolecules. The most important building blocks are (1) purines, in which the C-2 and C-8 atoms are introduced in folate-dependent reactions; (2) deoxythymidylate monophosphate (dTMP), synthesized from $N_{5,10}$-methylenetetrahydrofolate and deoxyuridylate monophosphate (dUMP); and (3) methionine, formed by the transfer of a methyl group from N_5-methyltetrahydrofolate to homocysteine (two of these three reactions are shown inFig. 107-1).

In all but one of the 1-carbon transfer reactions, tetrahydrofolate is produced. It can immediately accept a 1-carbon moiety and reenter the donor pool. The single exception is the thymidylate synthase reaction (dUMP®dTMP), in which dihydrofolate is the product (Fig. 107-1). This must be reduced to tetrahydrofolate by the enzyme dihydrofolate reductase before it can reenter the donor pool. A number of drugs are able to inhibit dihydrofolate reductase (Table 107-1), thereby diverting folate from the donor pool and producing what amounts to a state of folate deficiency in the face of normal tissue folate concentrations.

COBALAMIN

This vitamin is a complex organometallic compound in which a cobalt atom is situated within a corrin ring, a structure similar to the porphyrin from which heme is formed. Unlike heme, however, cobalamin cannot be synthesized in the human body and must be supplied in the diet. The only dietary source of cobalamin is animal products: meat and dairy foods. The minimum daily requirement for cobalamin is about 2.5 ug.

During gastric digestion, cobalamin in food is released and forms a stable complex with gastric R binder, one of a closely related group of glycoproteins of unknown function that are found in secretions (e.g., saliva, milk, gastric juice, bile), phagocytes, and plasma. On entering the duodenum, the cobalamin-R binder complex is digested, releasing the cobalamin, which then binds to intrinsic factor (IF), a 50-kDa glycoprotein produced by the parietal cells of the stomach. The secretion of IF generally parallels that of hydrochloric acid. The cobalamin-IF complex is resistant to proteolytic digestion and travels to the distal ileum, where specific receptors on the mucosal brush border bind and absorb the cobalamin-IF complex. Thus IF, like transferrin for iron, serves as a cell-directed carrier protein. The receptor-bound cobalamin-IF complex is taken into the ileal mucosal cell, where the IF is destroyed and the cobalamin is transferred to another transport protein, transcobalamin (TC) II. The cobalamin-TC II complex is then secreted into the circulation, from which it is rapidly taken up by the liver, bone marrow, and other cells. The pathway of cobalamin absorption is shown inFig. 107-2. Normally, about 2 mg cobalamin is stored in the liver, and another 2 mg is stored elsewhere in the body. In view of the minimum daily requirement, about 3 to 6 years would be required for a normal individual to become deficient in cobalamin if absorption were to cease abruptly.

AlthoughTC II is the acceptor for newly absorbed cobalamin, most circulating cobalamin is bound to TC I, a glycoprotein closely related to gastric R binder. TC I appears to be derived in part from leukocytes. The paradox that most circulating cobalamin is bound to TC I rather than TC II, even though TC II initially carries all the cobalamin that is absorbed by the intestine, is explained by the fact that cobalamin bound to TC II is rapidly cleared from the blood ($t_{1/2}$about 1 h), while clearance of cobalamin bound to TC I requires many days. The function of TC I is unknown.

Cobalamin is an essential cofactor for two enzymes in human cells: methionine synthase and methylmalonyl-CoA synthase. Cobalamin exists in two metabolically active forms, identified by the alkyl group attached to the sixth coordination position of the cobalt atom: methylcobalamin and adenosylcobalamin. The vitamin preparation that is used therapeutically is cyanocobalamin (also called vitamin B_{12}). Cyanocobalamin has no known physiologic role and must be converted to a biologically active form before it can be used by tissues.

Methylcobalamin is the form required for methionine synthase, which catalyzes the conversion of homocysteine to methionine (Fig. 107-1). When this reaction is impaired, folate metabolism is deranged, and it is this derangement that underlies the defect in DNA synthesis and the megaloblastic maturation pattern in patients who are deficient in cobalamin. In cobalamin deficiency, the unconjugated N_5-methyltetrahydrofolate newly taken from the bloodstream cannot be converted to other forms of tetrahydrofolate by methyl transfer. This is the so-called folate trap hypothesis. Because N_5-methyltetrahydrofolate is a poor substrate for the conjugating enzyme, it largely remains in the unconjugated form and slowly leaks from the cell. Tissue folate deficiency therefore develops, and this results in megaloblastic hematopoiesis. This hypothesis explains why tissue folate stores in cobalamin deficiency are substantially reduced, with a disproportionate reduction in conjugated, as compared with unconjugated, folates, despite normal or supranormal serum folate levels. It also explains why large doses of folate can produce a partial hematologic remission in patients with cobalamin deficiency.

Megaloblastic changes in both cobalamin and folate deficiency as well as in methotrexate treatment are related to a deficiency in production ofdTMP. In addition, the excess deoxyuridylate that accumulates can be phosphorylated and mistakenly incorporated into DNA in place of thymidylate; base pairing can be affected by this U-for-T substitution.

Plasma homocysteine levels are elevated in both folate and cobalamin deficiency, and high levels of plasma homocysteine appear to be a risk factor for thrombosis in both veins and arteries. It is not yet known, however, if hyperhomocysteinemia due to folate or cobalamin deficiency predisposes to thrombosis or alters its response to treatment.

Impairment in the conversion of homocysteine to methionine may also be partly responsible for the neurologic complications of cobalamin deficiency (see below). The methionine formed in this reaction is needed for the production of choline and choline-containing phospholipids. Nervous system damage is postulated to result at least in part from interference with these processes due to decreased methionine

production in cobalamin deficiency.

Adenosylcobalamin is required for the conversion of methylmalonyl CoA to succinyl CoA. Lack of this cofactor leads to large increases in the tissue levels of methylmalonyl CoA and its precursor, propionyl CoA. As a consequence, nonphysiologic fatty acids containing an odd number of carbon atoms are synthesized and incorporated into neuronal lipids. This biochemical abnormality may also contribute to the neurologic complications of cobalamin deficiency (see below).

CLINICAL DISORDERS

CLASSIFICATION OF MEGALOBLASTIC ANEMIAS (Table 107-1)

The cause of megaloblastic anemia varies in different parts of the world. In temperate zones, folate deficiency in alcoholics and pernicious anemia are the common types of megaloblastic anemias. In certain areas close to the equator, tropical sprue is endemic and an important cause of megaloblastic anemia, while in Scandinavia, infestations by the fish tapeworm, *Diphyllobothrium latum*, may be a cause.

The dietary intake of cobalamin is more than adequate for the body's requirements, except in true vegetarians and their breast-fed infants. Thus deficiency of cobalamin is almost always due to malabsorption. Malabsorption can occur at several levels. In contrast, the dietary intake of folic acid is marginal in many parts of the world. Furthermore, because the body's stores of folate are relatively low, folic acid deficiency can arise rather suddenly during periods of decreased dietary intake or increased metabolic demand. Finally, folic acid deficiency may be due to malabsorption. Often two or more of these factors coexist in a given patient.

Combined deficiencies of cobalamin and folic acid are not uncommon. Patients with tropical sprue are often deficient in both vitamins. The biochemical lesion that results in megaloblastic maturation of bone marrow cells also causes structural and functional abnormalities of the rapidly proliferating epithelial cells of the intestinal mucosa. Thus severe deficiency of one vitamin can lead to malabsorption of the other. Furthermore, as discussed above, a deficiency of cobalamin causes a secondary reduction in cellular folic acid.

Finally, megaloblastic anemias may occasionally be induced by factors unrelated to a vitamin deficiency. Most such cases are caused by one or more of the many drugs that interfere with DNA synthesis. Less commonly, megaloblastic maturation is encountered in certain acquired defects of hematopoietic stem cells. Rarest of all are specific congenital enzyme deficiencies.

COBALAMIN DEFICIENCY

The clinical features of cobalamin deficiency involve the blood, the gastrointestinal tract, and the nervous system.

The hematologic manifestations are almost entirely the result of anemia, although very rarely purpura may appear, due to thrombocytopenia. Symptoms of anemia may include

weakness, light-headedness, vertigo, and tinnitus, as well as palpitations, angina, and the symptoms of congestive failure. On physical examination, the patient with florid cobalamin deficiency is pale, with slightly icteric skin and eyes. Elevated bilirubin levels are related to high erythroid cell turnover in the marrow. The pulse is rapid, and the heart may be enlarged; auscultation will usually reveal a systolic flow murmur.

The gastrointestinal manifestations reflect the effect of cobalamin deficiency on the rapidly proliferating gastrointestinal epithelium. The patient sometimes complains of a sore tongue, which on inspection will be smooth and beefy red. Anorexia with moderate weight loss may also be evident, possibly accompanied by diarrhea and other gastrointestinal symptoms. These latter manifestations may be caused in part by megaloblastosis of the small intestinal epithelium, which results in malabsorption.

The neurologic manifestations often fail to remit fully on treatment. They begin pathologically with demyelination, followed by axonal degeneration and eventual neuronal death; the final stage, of course, is irreversible. Sites of involvement include peripheral nerves; the spinal cord, where the posterior and lateral columns undergo demyelination; and the cerebrum itself. Signs and symptoms include numbness and paresthesia in the extremities (the earliest neurologic manifestations), weakness, and ataxia. There may be sphincter disturbances. Reflexes may be diminished or increased. The Romberg and Babinski signs may be positive, and position and vibration senses are usually diminished. Disturbances of mentation will vary from mild irritability and forgetfulness to severe dementia or frank psychosis. It should be emphasized that *neurologic disease may occur in a patient with a normal hematocrit* and normalRBCindexes. Although it has many benefits, folate supplementation of food may increase the likelihood of neurologic presentations of cobalamin deficiency.

In the classic patient, in whom hematologic problems predominate, the blood and bone marrow show characteristic megaloblastic changes (described under "Diagnosis," below). The anemia may be very severe -- hematocrits of 15 to 20 are not infrequent -- but is surprisingly well tolerated by the patient because it develops so slowly.

Defective Release of Cobalamin from Food Cobalamin in food is tightly bound to enzymes in meat and is split from these enzymes by hydrochloric acid and pepsin in the stomach. People older than 70 years are commonly unable to release cobalamin from food sources but retain the ability to absorb crystalline B_{12}, the form most commonly found in multivitamins. The exact incidence of the defect in cobalamin release from food has not been well defined; estimates vary from 10 to greater than 50% of those over age 70 years. Only a minority of these persons go on to develop frank cobalamin deficiency, but many have biochemical changes, including low levels of cobalamin bound toTC II and elevated homocysteine levels, that augur cobalamin deficiency (see below).

Similarly, patients on drugs that suppress gastric acid production, such as omeprazole, may also fail to release cobalamin from food.

Pernicious Anemia Pernicious anemia, considered the most common cause of cobalamin deficiency, is caused by the absence ofIF, from either atrophy of the gastric mucosa or autoimmune destruction of parietal cells. It is most frequently seen in

individuals of northern European descent and African Americans and is much less common in southern Europeans and Asians. Men and women are equally affected. It is a disease of the elderly, the average patient presenting near age 60; it is rare under age 30, although typical pernicious anemia can be seen in children under age 10 (juvenile pernicious anemia). Inherited conditions in which a histologically normal stomach secretes either an abnormal IF or none at all will induce cobalamin deficiency in infancy or early childhood.

The incidence of pernicious anemia is substantially increased in patients with other diseases thought to be of immunologic origin, including Graves' disease, myxedema, thyroiditis, idiopathic adrenocortical insufficiency, vitiligo, and hypoparathyroidism. Patients with pernicious anemia also have abnormal circulating antibodies related to their disease: 90% have antiparietal cell antibody, which is directed against the H^+,K^+-ATPase, while 60% have anti-IF antibody. Antiparietal cell antibody is also found in 50% of patients with gastric atrophy without pernicious anemia, as well as in 10 to 15% of an unselected patient population, but anti-IF antibody is usually absent from these patients. Relatives of patients with pernicious anemia have an increased incidence of the disease, and even clinically unaffected relatives may have anti-IF antibody in their serum. Finally, treatment with glucocorticoids may reverse the disease.

The destruction of parietal cells in pernicious anemia is thought to be mediated by cytotoxic T cells. Pernicious anemia is unusually common in patients with agammaglobulinemia, suggesting that the cellular immune system plays a role in its pathogenesis. In contrast, *Helicobacter pylori* does not cause parietal cell destruction in pernicious anemia.

The most characteristic finding in pernicious anemia is gastric atrophy affecting the acid- and pepsin-secreting portion of the stomach; the antrum is spared. Other pathologic changes are secondary to the deficiency of cobalamin; these include megaloblastic alterations in the gastric and intestinal epithelium and the neurologic changes described above. The abnormalities in the gastric epithelium appear as cellular atypia in gastric cytology specimens, a finding that must be carefully distinguished from the cytologic abnormalities seen in gastric malignancy.

The *clinical manifestations* are primarily those of cobalamin deficiency, as described above. The disease is of insidious onset and progresses slowly. Laboratory examination will reveal hypergastrinemia and pentagastrin-fast achlorhydria as well as the hematologic and other laboratory abnormalities discussed under "Diagnosis."

Through appropriate replacement therapy, patients with pernicious anemia should experience complete and lifelong correction of all abnormalities that are due to cobalamin deficiency, except to the extent that irreversible changes in the nervous system may have occurred before treatment. These patients, however, are unusually subject to gastric polyps and have about twice the normal incidence of cancer of the stomach. Thus, patients should be followed with frequent stool guaiac examinations and endoscopy when indicated.

Postgastrectomy Following total gastrectomy or extensive damage to gastric mucosa as, for example, by ingestion of corrosive agents, megaloblastic anemia will develop

because the source ofIF has been removed. In all such patients, the absorption of orally administered cobalamin is impaired. Megaloblastic anemia may also follow partial gastrectomy, but the incidence is lower than after total gastrectomy. The cause of cobalamin deficiency after partial gastrectomy is not clear; defective release of cobalamin from food and intestinal overgrowth of bacteria have been suggested, but response to antibiotics is not common.

Intestinal Organisms Megaloblastic anemia may occur with intestinal stasis due to anatomic lesions (strictures, diverticula, anastomoses, "blind loops") or pseudoobstruction (diabetes mellitus, scleroderma, amyloid). This anemia is caused by colonization of the small intestine by large masses of bacteria that consume intestinal cobalamin before absorption. Steatorrhea may also be seen under these circumstances because bile salt metabolism is disturbed when the intestine is heavily colonized with bacteria. Hematologic responses have been observed after administration of oral antibiotics such as tetracycline and ampicillin. Megaloblastic anemia is seen in persons harboring the fish tapeworm, *D. latum*, due to competition by the worm for cobalamin. Destruction of the worm eliminates the problem.

Ileal Abnormalities Cobalamin deficiency is common in tropical sprue, while it is an unusual complication of nontropical sprue (gluten-sensitive enteropathy;Chap. 286). Virtually any disorder that compromises the absorptive capacity of the distal ileum can result in cobalamin deficiency. Specific entities include regional enteritis, Whipple's disease, and tuberculosis. Segmental involvement of the distal ileum by disease can cause megaloblastic anemia without any other manifestations of intestinal malabsorption such as steatorrhea. Cobalamin malabsorption is also seen after ileal resection. The Zollinger-Ellison syndrome (intense gastric hyperacidity due to a gastrin-secreting tumor) may cause cobalamin malabsorption by acidifying the small intestine, retarding the transfer of the vitamin from R binder toIF and impairing the binding of the cobalamin-IF complex to the ileal receptors. Chronic pancreatitis may also cause cobalamin malabsorption by impairing the transfer of the vitamin from R binder to IF. This abnormality can be detected by tests of cobalamin absorption (see below, Schilling test), but it is invariably mild and never causes clinical cobalamin deficiency. Finally, there is a rare congenital disorder, Imerslund-Grasbeck disease, in which a selective defect in cobalamin absorption is accompanied by proteinuria. Affected individuals have a mutation in cubulin, a receptor that mediates intestinal absorption of the cobalamin-IF complex.

Nitrous Oxide Inhalation of nitrous oxide as an anesthetic destroys endogenous cobalamin. As ordinarily used, the magnitude of the effects are not sufficient to cause clinical cobalamin deficiency, but repeated or protracted exposure (>6 h), particularly in older patients with borderline cobalamin stores, can lead to severe megaloblastic anemia and/or acute neurologic deficits.

FOLIC ACID DEFICIENCY

Since January, 1998, folic acid has been added to all enriched grain products by order of the U.S. Food and Drug Administration; accordingly, the incidence of folic acid deficiency has fallen markedly. Patients with folic acid deficiency are more often malnourished than those with cobalamin deficiency. The gastrointestinal manifestations

are similar to but may be more widespread and more severe than those of pernicious anemia. Diarrhea is often present, and cheilosis and glossitis are also encountered. However, in contrast to cobalamin deficiency, neurologic abnormalities do not occur.

The hematologic manifestations of folic acid deficiency are the same as those of cobalamin deficiency. Folic acid deficiency can generally be attributed to one or more of the following factors: inadequate intake, increased demand, or malabsorption.

Inadequate Intake Alcoholics may become folate deficient because their main source of caloric intake is alcoholic beverages. Distilled spirits are virtually devoid of folic acid, while beer and wine do not contain enough of the vitamin to satisfy the daily requirement. In addition, alcohol may interfere with folate metabolism. Narcotic addicts are also prone to become folate deficient because of malnutrition. Many indigent and elderly individuals who subsist primarily on canned foods or "tea and toast" and occasional teenagers whose diet consists of "junk food" develop folate deficiency. Food folate supplementation has made folate deficiency very rare.

Increased Demand Tissues with a relatively high rate of cell division such as the bone marrow or gut mucosa have a large requirement for folate. Therefore, patients with chronic hemolytic anemias or other causes of very active erythropoiesis may become deficient. Pregnant women formerly were at risk to become deficient in folic acid because of the high demand of the developing fetus. Deficiency in the first weeks of pregnancy can cause neural tube defects in newborns. Often the pregnancy was not detected until the defect had developed; thus, provision of folate supplementation to women after they learned they were pregnant was ineffective. However, folate food supplementation has decreased neural tube defects by more than 50%. Folate deficiency may also occur during the growth spurts of infancy and adolescence. Patients on chronic hemodialysis may require supplementary folate to replace that lost in the dialysate.

Malabsorption Folic acid deficiency is a common accompaniment of tropical sprue. Both the gastrointestinal symptoms and malabsorption are improved by the administration of either folic acid or antibiotics by mouth. Patients with nontropical sprue (gluten-sensitive enteropathy) may also develop significant folic acid deficiency that parallels other parameters of malabsorption. Similarly, folate deficiency in alcoholics may be due in part to malabsorption. In addition, other primary small-bowel disorders are sometimes associated with folate deficiency (Chap. 286).

DRUGS

Next to deficiency of folate or cobalamin, the most common cause of megaloblastic anemia is drugs. Agents that cause megaloblastic anemia do so by interfering with DNA synthesis, either directly or by antagonizing the action of folate. They can be classified as follows:

1. *Direct inhibitors of DNA synthesis*. They include purine analogues (6-thioguanine, azathioprine, 6-mercaptopurine), pyrimidine analogues (5-fluorouracil, cytosine arabinoside), and other drugs that interfere with DNA synthesis by a variety of mechanisms (hydroxyurea, procarbazine). The antiviral agent zidovudine (AZT), used

for treating HIV, often causes severe megaloblastic anemia.

2. *Folate antagonists*. The most toxic of these is methotrexate, a powerful inhibitor of dihydrofolate reductase which is used in the treatment of certain malignancies. Much less toxic but still capable of inducing a megaloblastic anemia are several weak dihydrofolate reductase inhibitors used to treat a variety of nonmalignant conditions including pentamidine, trimethoprim, triamterene, and pyrimethamine.

3. *Others*. A number of drugs antagonize folate by mechanisms that are poorly understood but are thought to involve an effect on absorption of the vitamin by the intestine. In this category are the anticonvulsants phenytoin, primidone, and phenobarbital. Megaloblastic anemia induced by these agents is mild.

OTHER MECHANISMS

Hereditary Megaloblastic anemia may be seen in several hereditary disorders. Orotic aciduria is a deficiency of orotidylic decarboxylase and phosphorylase, leading to a defect in pyrimidine metabolism and characterized by retarded growth and development as well as by the excretion of large amounts of orotic acid. Megaloblastic anemia has been reported in a single case of the Lesch-Nyhan syndrome, a condition resulting from a deficiency of hypoxanthine-guanine phosphoribosyltransferase whose clinical manifestations include gout, mental retardation, and self-mutilation. It has also been described in methylmalonic aciduria due to a combined defect in the biosynthesis of methyl and adenosyl cobalamins, although it is not seen in methylmalonic aciduria due to methylmalonyl CoA mutase deficiency. Congenital folate malabsorption causes megaloblastic anemia, accompanied by ataxia and mental retardation. Megaloblastic anemia has been reported to accompany the congenital deficiency of two other folate-metabolizing enzymes: dihydrofolate reductase and N_5-methyltetrahydrofolate:homocysteine methyltransferase. These deficiencies are less well documented than is congenital folate malabsorption. A thiamine-responsive megaloblastic anemia accompanied by nerve deafness and diabetes mellitus has been reported in several children. Megaloblastic changes as well as multinuclearity of red blood cell precursors are seen in the marrow of certain patients with congenital dyserythropoietic anemia, a group of inherited disorders characterized by mild to moderate anemia and a benign course.

TCII deficiency, like the congenital abnormalities in cobalamin absorption described previously, causes pronounced deficiency in cobalamin in infancy or early childhood, with all the accompanying manifestations. Megaloblastic anemia is not seen in hereditary TC I deficiency.

Refractory Megaloblastic Anemia This is a form of myelodysplasia in which megaloblastic erythropoiesis may sometimes be seen. Megaloblastic changes are restricted to the RBCseries (see below). As with other forms of myelodysplasia, refractory megaloblastic anemia is associated with an increased incidence of acute leukemia.

Megaloblastic changes are seen in erythremic myelosis and acute erythroleukemia (di Guglielmo), whereRBCprecursors are prominently involved. Here, the marrow is

characterized by bizarre erythroid maturation, with multinuclearity and multipolar mitotic figures in the RBC precursors (Chap. 111).

MEGALOBLASTIC DISEASE WITHOUT ANEMIA

Megaloblastic disease is easily overlooked in nonanemic patients. It can present in one of two ways.

Acute Megaloblastic Anemia Occasionally, a full-blown megaloblastic state can develop over the course of just a few days. This is usually seen following nitrous oxide anesthesia but may occur in any patient with a serious illness requiring intensive care, especially a patient receiving multiple transfusions, dialysis, or total parenteral nutrition. An acute megaloblastic state can also be precipitated by the administration of a weak antifolate (e.g., trimethoprim) to a patient with marginal tissue folate stores.

The condition resembles an immune cytopenia, with a rapidly developing thrombocytopenia and/or leukopenia in the absence of anemia. The blood smear may be completely normal, but the marrow is floridly megaloblastic. Acute megaloblastic anemia responds rapidly to treatment with folate plus cobalamin in the usual therapeutic doses.

Cobalamin Deficiency without Anemia Cobalamin deficiency without hematologic abnormalities is surprisingly common, especially in the elderly. The risk of a nonhematologic presentation for cobalamin deficiency is increased by the folate food fortification because folate can mask the hematologic effects of cobalamin deficiency. Between 10 and 30% of persons over age 70 years have metabolic evidence of cobalamin deficiency, either elevated homocysteine levels, low cobalamin-TCII levels, or both. Only 10% of these patients have defective production ofIF, and the remainder often have atrophic gastritis and cannot release cobalamin from their food (see above). These patients may present with neuropsychiatric abnormalities, including peripheral neuropathies, gait disturbance, memory loss, and psychiatric symptoms, sometimes with abnormal evoked potentials. Serum cobalamin levels may be normal or low, but serum levels of methylmalonic acid are almost invariably increased due to a deficiency of cobalamin at the tissue level. The neuropsychiatric abnormalities tend to improve and serum methylmalonic acid levels generally return to normal after treatment with cobalamin. Neurologic defects do not always reverse with cobalamin supplementation.

DIAGNOSIS

The finding of significant macrocytosis [mean corpuscular volume (MCV)> 100 fL] suggests the presence of a megaloblastic anemia. Other causes of macrocytosis include hemolysis, liver disease, alcoholism, hypothyroidism, and aplastic anemia. If the macrocytosis is marked (MCV > 110 fL), the patient is much more likely to have a megaloblastic anemia. Macrocytosis is less marked with concurrent iron deficiency or thalassemia. The reticulocyte count is low, and the leukocyte and platelet count may also be decreased, particularly in severely anemic patients. The blood smear (see Plate V-24) demonstrates marked anisocytosis and poikilocytosis, together with macroovalocytes, which are large, oval, fully hemoglobinized erythrocytes typical of megaloblastic anemias. There is some basophilic stippling, and an occasional

nucleatedRBC may be seen. In the white blood cell series, the neutrophils show hypersegmentation of the nucleus (see Plate V-38). This is such a characteristic finding that a single cell with a nucleus of six lobes or more should raise the immediate suspicion of a megaloblastic anemia. A rare myelocyte may also be seen. Bizarre, misshapen platelets are also observed. The reticulocyte index is low. The bone marrow is hypercellular with a decreased myeloid/erythroid ratio and abundant stainable iron. RBC precursors are abnormally large and have nuclei that appear much less mature than would be expected from the development of the cytoplasm (nuclear-cytoplasmic asynchrony). The nuclear chromatin is more dispersed than expected, and it condenses in a peculiar fenestrated pattern that is very characteristic of megaloblastic erythropoiesis. Abnormal mitoses may be seen. Granulocyte precursors are also affected, many being larger than normal, including giant bands and metamyelocytes. Megakaryocytes are decreased and show abnormal morphology.

Megaloblastic anemias are characterized by ineffective erythropoiesis (Chap. 61). In a severely megaloblastic patient, as many as 90% of theRBCprecursors may be destroyed before they are released into the bloodstream, compared with 10 to 15% in normal individuals. Enhanced intramedullary destruction of erythroblasts results in an increase in unconjugated bilirubin and lactic acid dehydrogenase (isoenzyme 1) in plasma. Abnormalities in iron kinetics also attest to the presence of ineffective erythropoiesis, with increased iron turnover but low incorporation of labeled iron into circulating RBCs.

In evaluating a patient with megaloblastic anemia, it is important to determine whether there is a specific vitamin deficiency by measuring serum cobalamin and folate levels. The normal range of cobalamin in serum is 200 to 900 pg/mL; values<100 pg/mL indicate clinically significant deficiency. Measurements of cobalamin bound toTC II would be a more physiologic measure of cobalamin status, but such assays are not yet routinely available. The normal serum concentration of folic acid ranges from 6 to 20 ng/mL; values £4 ng/mL are generally considered to be diagnostic of folate deficiency. Unlike serum cobalamin, serum folate levels may reflect recent alterations in dietary intake. Measurement ofRBCfolate level provides useful information because it is not subject to short-term fluctuations in folate intake and is better than serum folate as an index of folate stores.

Once cobalamin deficiency has been established, its pathogenesis can be delineated by means of a Schilling test. A patient is given radioactive cobalamin by mouth, followed shortly thereafter by an intramuscular injection of unlabeled cobalamin. The proportion of the administered radioactivity excreted in the urine during the next 24 h provides an accurate measure of absorption of cobalamin, assuming that a complete urine sample has been collected. Because cobalamin deficiency is almost always due to malabsorption (Table 107-1), this first stage of the Schilling test should be abnormal (i.e., small amounts of radioactivity in the urine). The patient is then given labeled cobalamin bound toIF. Absorption of the vitamin will now approach normal if the patient has pernicious anemia or some other type of IF deficiency. If cobalamin absorption is still decreased, the patient may have bacterial overgrowth (blind loop syndrome) or ileal disease (including an ileal absorptive defect secondary to the cobalamin deficiency itself). Cobalamin malabsorption due to bacterial overgrowth can frequently be corrected by the administration of antibiotics. The Schilling test can provide equally reliable

information after the patient has had adequate therapy with parenteral cobalamin.

A normal Schilling test in a patient with documented cobalamin deficiency may indicate poor absorption of the vitamin when mixed with food. This can be established by repeating the Schilling test with radioactive cobalamin scrambled with an egg.

Serum methylmalonic acid and homocysteine levels are also useful in the diagnosis of megaloblastic anemias. Both are elevated in cobalamin deficiency, while elevated levels of homocysteine but not methylmalonic acid are seen in folate deficiency. These tests measure tissue vitamin stores and may demonstrate a deficiency even when the more traditional but less reliable folate and cobalamin levels are borderline or even normal. Patients (particularly older patients) without anemia and with normal serum cobalamin levels but elevated levels of serum methylmalonic acid may develop neuropsychiatric abnormalities. Treatment of patients with this "subtle" cobalamin deficiency will usually prevent further deterioration and may result in improvement.

TREATMENT

Cobalamin Deficiency Apart from specific therapy related to the underlying disorder (e.g., antibiotics for intestinal overgrowth with bacteria), the mainstay of treatment for cobalamin deficiency is replacement therapy. Because the defect is nearly always malabsorption, patients are generally given parenteral treatment, specifically in the form of intramuscular cyanocobalamin. Parenteral treatment begins with 1000 ug cobalamin per week for 8 weeks, followed by 1000 ug cyanocobalamin intramuscularly every month for the rest of the patient's life. However, cobalamin deficiency can also be managed very effectively by oral replacement therapy with 2 mg crystalline B_{12} per day.

The response to treatment is gratifying. Shortly after treatment is begun, and several days before a hematologic response is evident in the peripheral blood, the patient will experience an increase in strength and an improved sense of well-being. Marrow morphology begins to revert toward normal within a few hours after treatment is initiated. Reticulocytosis begins 4 to 5 days after therapy is started and peaks at about day 7 (Fig. 107-3), with subsequent remission of the anemia over the next several weeks. If a reticulocytosis does not occur, or if it is less brisk than expected from the level of the hematocrit, a search should be made for other factors contributing to the anemia (e.g., infection, coexisting iron and/or folate deficiency, or hypothyroidism). Hypokalemia and salt retention may occur early in the course of therapy. Thrombocytosis may also be seen.

In most cases, replacement therapy is all that is needed for the treatment of cobalamin deficiency. Occasionally, however, a patient with a severe anemia will have such a precarious cardiovascular status that emergency transfusion is necessary. This must be done with great care, because such patients may develop heart failure from fluid overload. Blood must be administered slowly in the form of packedRBCs, with very close observation. A small volume of packed RBCs will frequently be enough to ameliorate the acute cardiovascular problems. If necessary, blood may be administered by exchanging patient blood (mostly plasma) for packed cells.

With lifelong treatment, patients should experience no further manifestations of

cobalamin deficiency, although neurologic symptoms may not be fully corrected even by optimal therapy. The potential for late development of gastric carcinoma in pernicious anemia necessitates careful follow-up of the patient.

Folate, particularly in large doses, can correct the megaloblastic anemia of cobalamin deficiency without altering the neurologic abnormalities. The neurologic manifestations may even be aggravated by folate therapy. Cobalamin deficiency can thus be masked in patients who are taking large doses of folate. For this reason, a hematologic response to folate must never be used to rule out cobalamin deficiency in a given patient; cobalamin deficiency can be excluded only by appropriate laboratory evaluation.

In light of the high frequency of defective cobalamin absorption in older people and the possible increased risk that overt cobalamin deficiency will present with neurologic rather than hematologic symptoms (because of folate food fortification), some experts have recommended the use of 0.1 mg oral crystalline cobalamin prophylaxis daily in people over age 65 years.

Folate Deficiency As for cobalamin deficiency, folate deficiency is treated by replacement therapy. The usual dose of folate is 1 mg/d, by mouth, but higher doses (up to 5 mg/d) may be required for folate deficiency due to malabsorption. Parenteral folate is rarely necessary. The hematologic response is similar to that seen after replacement therapy for cobalamin deficiency, i.e., a brisk reticulocytosis after about 4 days, followed by correction of the anemia over the next 1 to 2 months. The duration of therapy depends on the basis of the deficiency state. Patients with a continuously increased requirement (such as patients with hemolytic anemia) or those with malabsorption or chronic malnutrition should continue to receive oral folic acid indefinitely. In addition, the patient should be encouraged to maintain an optimal diet containing adequate amounts of folate.

Other Causes of Megaloblastic Anemia Megaloblastic anemia due to drugs can be treated, if necessary, by reducing the dose of the drug or eliminating it altogether. The effects of folate antagonists that inhibit dihydrofolate reductase can be counteracted by folinic acid [5-formyl tetrahydrofolate (THF)] in a dose of 100 to 200 mg/d (Fig. 107-1), which circumvents the block in folate metabolism by providing a form of folate that can be converted to 5,10-methylene THF. For the megaloblastic forms of sideroblastic anemia, pyridoxine in pharmacologic doses (as high as 300 mg/d) should be tried. If this fails, pyridoxal phosphate may work, presumably in part by promoting the conversion of THF to 5,10-methylene THF. Simple supportive measures are all that appear to be in order for treatment of refractory megaloblastic anemia. Acute erythroleukemia (di Guglielmo's disease) is usually treated like other types of acute myeloid leukemia (Chap. 111).

(Bibliography omitted in Palm version)

Back to Table of Contents

108. HEMOLYTIC ANEMIAS AND ACUTE BLOOD LOSS - *H. Franklin Bunn*, *Wendell Rosse*

The loss of red cells either through hemorrhage or, less commonly, through premature destruction of the red cells (hemolysis) may cause anemia. Hemolysis or blood loss normally leads to an increase in red cell production, which is clinically manifested by an increase in reticulocytes.

HEMOLYTIC ANEMIAS

Red blood cells (RBC) normally survive 90 to 120 days in the circulation. The life span of RBC may be shortened in a number of disorders, often resulting in anemia if the bone marrow is not able to replenish adequately the prematurely destroyed RBC. The disorders associated with hemolytic anemias are generally identified by the abnormality that brings about the premature destruction of the RBC.

In all patients with hemolytic anemia, a careful history and physical examination provide important clues to the diagnosis. The patient may complain of fatigue and other symptoms of anemia (Chap. 61). Less commonly, jaundice and even red-brown urine (hemoglobinuria) are reported. A complete drug and toxin exposure history and the family history often provide crucial information. The physical examination may show jaundice of skin and mucosae. Splenomegaly is encountered in a variety of hemolytic anemias. A wide array of other historic and physical findings is associated with specific hemolytic anemias (see below).

Laboratory tests may be used initially to demonstrate the presence of hemolysis (Table 108-1) and define its cause. An elevated reticulocyte count in the patient with anemia is the most useful indicator of hemolysis, reflecting erythroid hyperplasia of the bone marrow; biopsy of the bone marrow is often unnecessary. Reticulocytes are also elevated in patients with active blood loss, those with myelophthisis, and those who are recovering from suppression of erythropoiesis (Chap. 61). The morphology of theRBC may provide evidence both of hemolysis and of its cause; the characteristic abnormalities and their associated causes and syndromes are listed inTable 108-2. While the findings on the peripheral blood smear alone are rarely pathognomonic, they may provide important clues to the presence of hemolysis and to diagnosis.

RBCmay be prematurely removed from the circulation by macrophages, particularly those of the spleen and liver (extravascular lysis), or, less commonly, by disruption of their membranes during their circulation (intravascular hemolysis). Both mechanisms result in increased heme catabolism and enhanced formation of unconjugated bilirubin, which is normally conjugated by the liver and excreted. The plasma level of unconjugated bilirubin may be high enough to produce readily apparent jaundice (detectable usually when serum bilirubin is>34 umol/L or 2 mg/dL). The unconjugated (indirect) bilirubin level can be further elevated by a commonly encountered defect in conjugation of bilirubin (Gilbert's syndrome) (Chap. 294). In patients with hemolysis, the level of unconjugated bilirubin never exceeds 70 to 85 umol/L (4 to 5 mg/dL), unless liver function is impaired.

In the absence of tissue damage in other organs, serum enzyme levels can be useful in

the diagnosis and monitoring of patients with hemolysis. Lactate dehydrogenase (LDH), particularly LDH-2, is elevated by acceleratedRBCdestruction. Serum AST (SGOT) may be somewhat elevated, whereas ALT (SGPT) is not.

Haptoglobin is ana globulin that is present in high concentration (~1.0 g/L) in the plasma (and serum). It binds specifically and tightly to the globin in hemoglobin. The hemoglobin-haptoglobin complex is cleared within minutes by the mononuclear phagocyte system. Thus patients with significant hemolysis, either intravascular or extravascular, have low or absent levels of serum haptoglobin. The fact that haptoglobin synthesis is decreased in patients with hepatocellular disease and increased in inflammatory states must be considered in the interpretation of serum haptoglobin.

Intravascular hemolysis (which is uncommon) results in the release of hemoglobin into the plasma. In these cases, plasma hemoglobin is increased in proportion to the degree of hemolysis. Plasma hemoglobin may be falsely elevated due to lysis ofRBC in vitro. If the haptoglobin-binding capacity of the plasma is exceeded, free hemoglobin passes through renal glomeruli. This filtered hemoglobin is reabsorbed by the proximal tubule, where it is catabolized in situ, and the heme iron is incorporated into storage proteins (ferritin and hemosiderin). The presence of hemosiderin in the urine, detected by staining the sediment with Prussian blue, indicates that a significant amount of circulating free hemoglobin has been filtered by the kidneys. Hemosiderin appears 3 to 4 days after the onset of hemoglobinuria and may persist for weeks after its cessation. When the absorptive capacity of the tubular cells is exceeded, hemoglobinuria ensues. Hemoglobinuria indicates severe intravascular hemolysis. Hemoglobinuria must be distinguished from hematuria (in which case RBC are seen on urine examination) and from myoglobin due to rhabdomyolysis; in all three cases, the urine is positive with the benzidine reaction, commonly used in analysis of urine. The distinction between hemoglobinuria and myoglobinuria can best be made by specific tests that exploit immunologic differences or differences in solubility. After centrifugation of an anticoagulated blood specimen, the plasma of patients with hemoglobinuria has a reddish-brown color, whereas that of patients with myoglobinuria is normal in color. Because of its higher molecular weight, hemoglobin has lower glomerular permeability than myoglobin and is less rapidly cleared by the kidneys.

CLASSIFICATION

The hemolytic anemias can be grouped in three different ways, shown inTable 108-3. The cause of acceleratedRBCdestruction can be regarded as (1) a molecular defect (hemoglobinopathy or enzymopathy) inside the red cell, (2) an abnormality in membrane structure and function, or (3) an environmental factor such as mechanical trauma or an autoantibody. In *intracorpuscular types* of hemolysis, the patient's RBC have an abnormally short life span in a normal recipient (with a compatible blood type), while compatible normal RBC survive normally in the patient. The opposite is true in *extracorpuscular types* of hemolysis. Finally, hemolytic disorders can be classified as either inherited or acquired.

INHERITED HEMOLYTIC ANEMIAS

The inherited hemolytic anemias are due to inborn defects in one of three main

components of red cells: the membrane, the enzymes, or hemoglobin. These defects are often known at the genomic level, but their identification still largely depends on their clinical and laboratory manifestations.

Red Cell Membrane Disorders These are usually readily detected by morphologic abnormalities of the RBC on the blood film. There are three types of inherited RBC membrane abnormalities: hereditary spherocytosis, hereditary elliptocytosis (including hereditary pyropoikilocytosis), and hereditary stomatocytosis.

Hereditary spherocytosis This condition is characterized by spherical RBC due to a molecular defect in one of the proteins in the cytoskeleton of the RBC membrane, leading to a loss of membrane and hence decreased ratio of surface area to volume and consequently spherocytosis. This disorder usually has an autosomal dominant inheritance pattern and an incidence of approximately 1:1000 to 1:4500. In ~20% of patients, the absence of hematologic abnormalities in family members suggests either autosomal recessive inheritance or a spontaneous mutation. The disorder is sometimes clinically apparent in early infancy but often escapes detection until adult life.

CLINICAL MANIFESTATIONS The major clinical features of hereditary spherocytosis are anemia, splenomegaly, and jaundice. The prominence of jaundice accounts for the disorder's prior designation as "congenital hemolytic jaundice" and is due to an increased concentration of unconjugated (indirect-reacting) bilirubin in plasma. Jaundice may be intermittent and tends to be less pronounced in early childhood. Because of the increased bile pigment production, pigmented gallstones are common, even in childhood. Compensatory erythroid hyperplasia of the bone marrow occurs, with the extension of red marrow into the midshafts of long bones and occasionally with extramedullary erythropoiesis, at times leading to the formation of paravertebral masses visible on chest x-ray. Because the bone marrow's capacity to increase erythropoiesis by six- to eightfold exceeds the usual rate of hemolysis, anemia is usually mild or moderate and may even be absent in an otherwise healthy individual. Compensation may be temporarily interrupted by episodes of relative erythroid hypoplasia precipitated by infections, particularly parvovirus, trauma, surgery, and pregnancy. Splenomegaly is very common. The hemolytic rate may increase transiently during systemic infections, which induce further splenic enlargement. Chronic leg ulcers, similar to those observed in sickle cell anemia, occur occasionally.

The characteristic erythrocyte abnormality is the spherocyte (Plate V-26). The mean corpuscular volume (MCV) is usually normal or slightly decreased, and the mean corpuscular hemoglobin concentration (MCHC) is increased to 350 to 400 g/L. Spheroidicity may be quantitatively assessed by measurement of the osmotic fragility of theRBC on exposure to hypoosmotic solutions causing a net influx of water (Fig. 108-1). Because spherocytes have a decreased surface area per unit volume, they are able to take in less water and hence lyse at a higher concentration of saline than normal cells. On microscopic examination, spherocytes are usually detected as small cells without central pallor. They will ordinarily not influence the osmotic fragility test unless they constitute more than 1 or 2% of the total cell population. The autohemolysis test, which measures the amount of spontaneous hemolysis occurring after 48 h of sterile incubation, is also useful.

PATHOGENESIS The molecular abnormality in hereditary spherocytosis primarily involves the proteins responsible for tethering the lipid bilayer to the underlying cytoskeletal network. Nearly all patients have a significant deficiency of spectrin, which is sometimes secondary to an inherited molecular defect in that protein. About 50% of patients have a defect in ankyrin, the protein that forms a bridge between protein 3 and spectrin (Fig. 108-2). Homozygotes who have a recessive inheritance pattern for ankyrin deficiency have more severe anemia than heterozygotes with the more common dominant form. About 25% of patients have a mutation of protein 3, resulting in a deficiency of that protein and mild anemia with dominant inheritance. Most of the remaining 25% have mutations of spectrin, leading to impaired synthesis or self-association;b-spectrin deficiency is generally mild, with dominant inheritance, while a-spectrin deficiency is severe, with a recessive inheritance pattern. Because the lipid bilayer is not well anchored when these proteins are defective, part of it is lost by vesiculation, resulting in a more spherical and less deformable cell. Because of their shape and rigidity, spherocytes cannot traverse the interstices of the spleen where their increased metabolic rate cannot be sustained, causing a further loss of surface membrane. This "conditioning" produces a subpopulation of hyperspheroidalRBC in the peripheral blood.

DIAGNOSIS Hereditary spherocytosis must be distinguished primarily from the spherocytic hemolytic anemias associated withRBCantibodies. The family history of anemia and/or splenectomy is helpful, when present. The diagnosis of immune spherocytosis is usually readily established by a positive direct Coombs test (see below). Spherocytes are also seen in association with hemolysis induced by splenomegaly in patients with cirrhosis, in clostridial infections, and in certain snake envenomations (due to the action of phospholipases on the membrane). A few spherocytes are seen in the course of a wide variety of hemolytic disorders, particularly glucose-6-phosphate dehydrogenase (G6PD) deficiency.

TREATMENT

Splenectomy reliably corrects the anemia, although theRBCdefect and its consequent morphology persist. The operative risk is low. RBC survival after splenectomy is normal or nearly so; if it is not, an accessory spleen or another diagnosis should be sought. Because of the potential for gallstones and for episodes of bone marrow hypoplasia or hemolytic crises, splenectomy should be performed in symptomatic individuals; cholecystectomy should not be performed without splenectomy, as intrahepatic gallstones may result. Splenectomy in children should be postponed until age 4, if possible, to minimize the risk of severe infections with gram-positive encapsulated organisms. Polyvalent pneumococcal vaccine should be administered at least 2 weeks before splenectomy. In patients with severe hemolysis, folic acid (1 mg/d) should be administered prophylactically.

Hereditary elliptocytosis and hereditary pyropoikilocytosis Oval or ellipticRBC are normally found in birds, reptiles, camels, and llamas; however, they occur in appreciable numbers in humans only in *hereditary elliptocytosis*, a disorder that is transmitted as an autosomal dominant trait and affects 1 per 4000 to 5000 people, a frequency similar to that of hereditary spherocytosis (rarely, patients with myelodysplastic disorders of the bone marrow may have acquired elliptocytosis). The elliptic shape is acquired as the

cell deforms to traverse the microcirculation but does not spring back to its initial biconcave shape. In most affected individuals, a structural abnormality of erythrocyte spectrin that leads to impaired assembly of the cytoskeleton. In some families, affected individuals have a deficiency of erythrocyte membrane protein 4.1, which stabilizes the interaction of spectrin and actin in the cytoskeleton (Fig. 108-2); homozygotes with absence of this protein have more marked hemolysis. In Southeast Asia, there is a high incidence of hereditary ovalocytosis, in which a small internal deletion of protein 3 makes the membrane rigid and confers resistance against malaria.

The great majority of patients manifest only mild hemolysis, with hemoglobin levels>120 g/L, reticulocytes<4% (0.2 ´10_{12}/L), depressed haptoglobin levels, andRBCsurvival times just under the normal range. In 10 to 15% of patients with more severe abnormalities, the rate of hemolysis is substantially increased, with median survival times of RBC as short as 5 days and reticulocytes ranging up to 20%. Hemoglobin levels rarely fall below 90 to 100 g/L. RBC destruction occurs predominantly in the spleen, which is enlarged in patients with overt hemolysis. Hemolysis is corrected by splenectomy.

In both the anemic and nonanemic varieties of this disorder, at least 25% and, more commonly, >75% ofRBC are elliptic, with an axial ratio (width/length) of<0.78. Patients with hemolysis frequently have microovalocytes, bizarre-shaped RBC, and RBC fragments, all of which increase in number after splenectomy. The degree of hemolysis does not correlate with the percentage of elliptocytes. Osmotic fragility is usually normal but may be increased in patients with overt hemolysis.

Hereditary pyropoikilocytosis is a rare disorder related to hereditary elliptocytosis and is characterized by bizarre-shaped, microcyticRBC that undergo disruption at temperatures of 44 to 45°C (in contrast, normal RBC are stable up to 49°C). This condition results from a deficiency of spectrin and an abnormality of spectrin self-assembly. Hemolysis is usually severe, is recognized in childhood, and is partially responsive to splenectomy.

Hereditary Stomatocytosis Stomatocytes are cup-shapedRBC(concave on one face and convex on the other). This formation results in a slitlike central zone of pallor on dried smears. The syndrome of hereditary hemolytic anemia and stomatocytic RBC is inherited in an autosomal dominant pattern. RBC have an increased permeability to sodium and potassium, which is compensated for by an increased active transport of these cations. In some patients, the RBC are swollen with an excess of ions and water and a decreased mean corpuscular hemoglobin concentration (overhydrated stomatocytes, "hydrocytosis"); many of these patients lack the RBC membrane protein 7.2 (stomatin). In other patients, the RBC are shrunken, with a decreased ion and water content and an increased mean corpuscular hemoglobin concentration (dehydrated stomatocytes, "desiccytosis" or "xerocytosis"). Those patients in whom the RBC are overhydrated have true stomatocytes on dried smears. Dehydrated stomatocytes assume the morphology of target cells on dried smears. Osmotic fragility is increased in overhydrated stomatocytes and decreased in underhydrated stomatocytes. RBC lacking Rh proteins (Rh_{null}cells) are stomatocytic and have a shortened life span.

Most patients have splenomegaly and mild anemia. Splenectomy decreases but does

not totally correct the hemolytic process.

Red Cell Enzyme Defects During its maturation, theRBCloses its nucleus, ribosomes, and mitochondria and thus its capability for protein synthesis and oxidative phosphorylation. The mature circulating RBC has a relatively simple pattern of intermediary metabolism (Fig. 108-3) in keeping with its modest metabolic obligations. ATP must be generated from the Embden-Meyerhof pathway to drive the cation pump that maintains the ionic milieu in the RBC. Smaller amounts of energy are needed for the preservation of hemoglobin iron in the ferrous (Fe^{2+}) state and perhaps for the renewal of the lipids in the RBC membrane. About 10% of the glucose consumed by the RBC is metabolized via the hexose-monophosphate shunt (Fig. 108-3), which protects both hemoglobin and the membrane from exogenous oxidants, including certain drugs.

Deficiency states have been reported for most of the enzymes shown in Fig. 108-3. Many of these enzyme abnormalities appear to be restricted toRBC. Mutations can result in no protein product, a dysfunctional product, or an unstable product. Mutations resulting in decreased stability will be detected more readily in RBC compared with other cells having either a shorter life span or the synthetic capability to renew the enzyme.

Defects in the Embden-Meyerhof Pathway In general, these enzymopathies have similar pathophysiologic and clinical features. Patients present with a congenital nonspherocytic hemolytic anemia of variable severity. TheRBC are often relatively deficient in ATP, resulting in a leak of potassium ion out of these cells. Abnormalities in RBC morphology (see below) indicate that the membrane is affected by the enzyme defect. These RBC are rigid and thus more readily sequestered by the mononuclear phagocyte system.

Some of these glycolytic enzyme deficiencies such as pyruvate kinase (PK) deficiency and hexokinase deficiency are localized to theRBC, with no apparent metabolic abnormality in other cells; in the case of PK deficiency, this is due to specific isozymes confined to the RBC. In other disorders, the enzyme deficiency is more widespread. Glucose phosphate isomerase deficiency and phosphoglycerate kinase deficiency also involve leukocytes, although affected individuals have no apparent abnormalities of leukocyte function. Individuals with deficiency of triose phosphate isomerase have decreased levels of enzyme in leukocytes, muscle cells, and cerebrospinal fluid, and they have a progressive neurologic disorder. Some patients with phosphofructokinase deficiency have a myopathy.

About 95% of the clinically significant defects in the glycolytic pathway are due toPKdeficiency, and about 4% are due to glucose phosphate isomerase deficiency. The remainder, shown in Fig. 108-3, are extremely rare. Most have been encountered in isolated families; clinical manifestations are variable. A number of different mutations result in PK deficiency. Some missense mutations in decreased reaction with substrate (phosphoenolpyruvate), an enhancing molecule (fructose 1,6 diphosphate), or ADP. Thus, there is considerable variability in the clinical manifestations and laboratory findings among individuals reported as having PK deficiency. Most of these patients are compound heterozygotes who have inherited a different defective enzyme from each parent.

Most of the glycolytic enzyme defects are inherited in an autosomal recessive pattern. The parents of affected patients are heterozygotes and express half-normal levels of enzyme activity, which are more than adequate for normal metabolic function. Thus, the parents are entirely asymptomatic. Since the gene frequency for this group of enzymopathies is low, true homozygotes are often the offspring of a consanguineous mating. More often, affected individuals are compound heterozygotes. Phosphoglycerate kinase deficiency is inherited as a sex-linked disorder. Affected males have a severe hemolytic anemia, while female carriers may have a mild hemolytic process.

CLINICAL MANIFESTATIONS Patients with severe hemolysis usually present during early childhood with anemia, jaundice, and splenomegaly.

LABORATORY FINDINGS Patients have a normocytic (or slightly macrocytic), normochromic anemia with reticulocytosis. In those withPKdeficiency, bizarre erythrocytes, including spiculated cells, are noted on the peripheral smear, especially after splenectomy. Spherocytes are usually absent; hence the term *congenital nonspherocytic hemolytic anemia* has been applied to these disorders. Unlike hereditary spherocytosis, the osmotic fragility of freshly drawn blood is usually normal. Incubation brings out an osmotically fragile population ofRBC, an abnormality not corrected by the addition of glucose.

The diagnosis of this group of anemias depends on specific enzymatic assays; care must be taken to provide an appropriate concentration of substrate to detect those variants with a low affinity for substrate or enhancing molecule. An abnormality in enzyme kinetics, differences in electrophoretic mobility, pH optimum, or heat stability may be useful in documenting heterogeneity among enzyme variants.

TREATMENT

Most patients do not require therapy. Those with severe hemolysis should be given folic acid (1 mg/d). Blood transfusions may be necessary during a hypoplastic crisis. Women withPKdeficiency may become very anemic during pregnancy, sometimes leading to the diagnosis for the first time.

Because of their enzymatic defect, the younger cells (reticulocytes) depend on mitochondrial respiration rather than glycolysis for maintenance of ATP. However, in the hypoxic environment of the spleen, aerobic metabolism is curtailed and the ATP-depleted cells are destroyed in situ. Reticulocytes are normally retained in the spleen for 24 to 48 h. Patients withPKdeficiency may benefit from splenectomy, which usually leads to a marked increase in circulating reticulocytes. Patients with deficiency of glucose phosphate isomerase also may improve after splenectomy. Splenectomy has not been proven effective in individuals with other glycolytic enzymopathies.

Defects in the hexose-monophosphate shunt The normalRBC is well protected against oxidant stress. When the cell is exposed to a drug or toxin that generates oxygen radicals, glucose metabolism via the hexose-monophosphate shunt is normally increased severalfold. Reduced glutathione is regenerated, protecting the sulfhydryl

groups of hemoglobin and the RBC membrane from oxidation. Individuals with an inherited defect in the hexose-monophosphate shunt are unable to maintain an adequate level of reduced glutathione in their RBC, hemoglobin sulfhydryl groups become oxidized, and the hemoglobin precipitates within the RBC, forming Heinz bodies.

G6PD DEFICIENCY This is by far the most common congenital shunt defect, affecting more than 200 million people throughout the world; like hemoglobin S, it partially protects the patient from malaria by providing a defective home for the merozoite. Considerable genetic heterogeneity exists among affected individuals, and over 400 variants ofG6PDhave been described. In most cases, the alteration is a base substitution, leading to an amino acid replacement rather than a deletion or truncation of the protein. The mutations generate enzymes with differences in electrophoretic mobility, enzyme kinetics, pH optimum, and heat stability. These differences result in great variation of clinical severity, ranging from nonspherocytic hemolytic anemia without demonstrable oxidant stress (particularly shortly after birth), through hemolytic anemia only when stimulated by marked to mild oxidant stress, to no clinically detectable abnormality. The normal G6PD is designated as type B. About 20% of individuals of African descent have a G6PD (designated A+) that differs by a single amino acid and is electrophoretically distinguishable but functionally normal. Among the clinically significant G6PD variants, the most common, the so-called A- type, is due to two base substitutions and is encountered primarily in individuals of central African descent. The A- G6PD has the same electrophoretic mobility as the A+ type, but it is unstable and has abnormal kinetic properties. This variant is found in about 11% of African American males. A second relatively common G6PD variant is encountered among peoples of Mediterranean origin, particularly Sardinians and Sephardic Jews; this variant is more severe than the A- variant and may result in nonspherocytic hemolytic anemia in the absence of known oxidative stress. A third relatively common and slightly less severe variant occurs in southern Chinese populations.

TheG6PDgene is located on the X chromosome; thus the deficiency state is a sex-linked trait. Affected males (hemizygotes) inherit the abnormal gene from their mothers who are usually carriers (heterozygotes). Because of inactivation of one of the two X chromosomes (Lyon hypothesis:Chap. 65), the heterozygote has two populations ofRBC: normal and deficient in G6PD. Most female carriers are asymptomatic. Those who happen to have a high proportion of deficient cells resemble the male hemizygotes. G6PD activity normally declines ~50% during the 120-day life span of the RBC. This decay is moderately accelerated in A-RBC and markedly so in RBC containing the Mediterranean variant. Individuals with the A- variant normally have a slightly shortened RBC survival time, but they are not anemic. Clinical problems arise only when the affected individual is subjected to some type of environmental stress. Most often, hemolytic episodes are triggered by viral and bacterial infections. The mechanism is unknown. In addition, drugs or toxins that pose an oxidant threat to the RBC (most commonly sulfa drugs, antimalarials, and nitrofurantoin) cause hemolysis in individuals deficient in G6PD (Table 108-4). Although aspirin is frequently mentioned as a likely offender, it has no deleterious effect in A- individuals. Accidental ingestion of toxic compounds such as naphthalene (moth balls) may cause severe hemolysis. Metabolic acidosis can precipitate an episode of hemolysis in individuals deficient in G6PD.

CLINICAL AND LABORATORY FEATURES The patient may experience an acute hemolytic crisis within hours of exposure to the oxidant stress, leading to hemoglobinuria and peripheral vascular collapse in severe cases. Since only the older population ofRBC is rapidly destroyed, the hemolytic crisis is usually self-limited, even if the exposure to the oxidant continues. Among black males with the A- variant, the RBC mass decreases by a maximum of 25 to 30%. During acute hemolysis, a rapid drop in hematocrit is accompanied by a rise in plasma hemoglobin and unconjugated bilirubin and a decrease in plasma haptoglobin. The oxidation of hemoglobin leads to the formation of Heinz bodies, visualized by means of a supravital stain such as crystal violet. However, Heinz bodies are usually not seen after the first day or so, since these inclusions are readily removed by the spleen. Their removal leads to the formation of "bite cells" (RBC that have lost a peripheral portion of the cell). Multiple bites cause the formation of fragments. A few spherocytes also may be present. Individuals with the Mediterranean typeG6PDhave a more unstable enzyme and, therefore, a much lower overall enzyme activity than individuals with the A- variant. As a result, they have more severe clinical manifestations. A minority of patients are exquisitely sensitive to fava beans and develop a fulminant hemolytic crisis after exposure. The oxidants in *Vicia fava* are twob-glycosides whose aglycones, when autooxidized, produce oxygen free radicals. The incidence of favism is highly variable due to variations in concentration, in absorption, or in metabolism of the aglycones. Favism is not encountered in individuals with the A-variant.

The *diagnosis* ofG6PDdeficiency should be considered in any individual, particularly a male of African or Mediterranean descent, who experiences an acute hemolytic episode. The patient should be questioned about possible exposure to oxidant agents. The diagnosis can be established by a number of tests that assess either the enzyme activity or the effects of its deficiency. However, the test may yield a false-negative result during a hemolytic episode when the oldRBCcontaining the defective enzyme have already lysed.

TREATMENT

Since hemolysis in patients deficient in A-G6PD is usually self-limited, no specific treatment is necessary. Splenectomy does not benefit Mediterranean patients with chronic hemolysis. Blood transfusions are rarely indicated. Adequate urine flow should be maintained if hemoglobinuria develops during an acute hemolytic episode.

Prevention of hemolytic episodes is best. Infections ought to be treated promptly. Patients should be warned about risks posed by oxidant drugs and fava beans. Any patient of African or Mediterranean ancestry about to be given an oxidant drug should be screened forG6PDdeficiency.

OTHER DEFECTS OF THE HEXOSE-MONOPHOSPHATE SHUNT A few kindreds have been found to have congenital deficiency inRBCglutathione due to a defect in either of the two enzymes responsible for the synthesis of this tripeptide. Affected individuals have a hemolytic anemia with Heinz bodies that is aggravated by oxidant drugs. Deficiency of glutathione reductase has been reported, but its relationship to clinically significant hemolysis is not well established. Sometimes the deficiency state can be corrected by the administration of riboflavin (5 mg/d). Deficiencies of glutathione

peroxidase and 6-phosphogluconate dehydrogenase have been observed, but their association with hemolysis is uncertain.

Other enzyme defects Hemolytic anemia may sometimes be caused by abnormalities in enzymes of nucleotide metabolism. Individuals with pyrimidine 5¢-nucleotidase deficiency have marked coarse basophilic stippling in theirRBCbecause the mRNA of the cell is not properly metabolized. Hemolytic anemia also has been noted in individuals whose RBC have supranormal levels of adenosine deaminase and relatively low levels of ATP.

Hemoglobinopathies The sickling disorders constitute an important form of congenital hemolytic anemia. Less commonly, hemolysis may be due to the inheritance of an unstable hemoglobin variant.**For further discussion, see Chap. 106.*

ACQUIRED HEMOLYTIC ANEMIAS

In most patients with acquired hemolytic anemia,RBC are made normally but are prematurely destroyed because of damage acquired in the circulation. (The exceptions are rare disorders characterized by acquired dysplasia of the cells of the bone marrow and the production of structurally and functionally abnormal RBC.) The damage that occurs may be mediated by antibodies or toxins or may be due to abnormalities in the circulation, including an overactive mononuclear phagocyte system or traumatic lysis by natural or artificial impediments to circulation. The acquired hemolytic anemias can be classified into five categories (Table 108-5).

Hypersplenism The spleen is particularly efficient in trapping and destroyingRBC that have minimal defects. This unique ability of the spleen to filter mildly damaged RBC results from its unusual vascular anatomy (Chap. 63). Almost all the blood circulating through the spleen flows rapidly from arterioles in the white pulp to sinuses in the spleen's red pulp and then into the venous system. In contrast, a small portion of splenic blood flow (normally 1 to 2%) passes into the "marginal zone" of the lymphatic white pulp. Although the cells that occupy this zone are not phagocytic, they serve as a mechanical filter that hinders the progress of severely damaged blood cells. As RBC leave this zone and enter the red pulp, they flow into narrow cords, rich in macrophages, that end blindly but communicate with sinuses through small openings between the lining cells of the sinuses. These openings, averaging 3 um in diameter, test the ability of RBC (4.5 um in diameter) to undergo a deformation. RBC that cannot re-enter the vascular sinuses are engulfed by phagocytic cells and destroyed (seeFig. 63-1).

The normal spleen retains reticulocytes for 1 to 2 days but otherwise poses no threat to normalRBCuntil they become senescent. However, in the face of splenomegaly, increased destruction of the cells of the blood, including the RBC, may take place due to pooling of the blood in a relatively nutrient-poor environment full of phagocytic cells. When splenic sequestration causes cytopenia, hypersplenism is diagnosed. In infiltrative diseases of the spleen, substantial splenomegaly may exist with no apparent hemolysis; inflammatory and congestive splenomegaly is commonly associated with modest shortening of RBC survival time, along with more marked granulocytopenia and thrombocytopenia. Patients with cytopenia(s) sufficient to produce symptoms generally benefit from splenectomy.

Immunologic Causes of Hemolysis Immune hemolysis in the adult is usually induced by IgG or IgM antibodies with specificity for antigens associated with the patient'sRBC(often called "autoantibodies") (Table 108-6); rarely, transfused RBC may be hemolyzed by alloantibodies directed against foreign antigens on those cells (Chap. 114).

The Coombs antiglobulin test is the major tool for diagnosing autoimmune hemolysis. This test relies on the ability of antibodies specific for immunoglobulins (especially IgG) or complement components (especially C3) to agglutinateRBC when these proteins are present on the RBC. The *direct Coombs test* measures the ability of anti-IgG or anti-C3 antisera to agglutinate the patient's RBC. The presence or absence of IgG and/or C3 may help define the origin of the immune hemolytic anemia (Table 108-6). Rarely, neither IgG nor complement may be found on the RBC of the patient (Coombs-negative immune hemolytic anemia).

Antibodies to particularRBCantigens in the serum of the patient can be detected by reacting the serum with normal RBC bearing the antigen. IgM antibodies (usually cold-reacting) may be detected by agglutination of normal or fetal RBC. IgG antibodies may be detected by the *indirect Coombs test*, in which the serum of the patient is incubated with normal RBC and antibody is detected with anti-IgG, as in the direct Coombs test.

"Warm" antibodies Antibodies that react with protein antigens are nearly always IgG and react at body temperature; occasionally, they are IgA and rarely IgM. Hemolysis due to autologous antibodies is called *autoimmune hemolytic (or immunohemolytic) anemia, warm antibody type*.

CLINICAL MANIFESTATIONS Immunohemolytic anemia of the warm antibody type is induced by IgG antibody and occurs at all ages, but it is more common in adults, particularly women. In approximately one-fourth of patients this disorder occurs as a complication of an underlying disease affecting the immune system, especially lymphoid neoplasms (Chap. 112); collagen vascular diseases, especially systemic lupus erythematosus (SLE); and congenital immunodeficiency diseases (Table 108-7). In the initial evaluation of the patient, drugs that are known to cause immunohemolytic anemia must be ruled out (see below). The presentation and course of IgG immunohemolytic anemia are quite variable. In its mildest form, the only manifestation is a positive direct Coombs test. In this instance, insufficient antibody is present on theRBCsurface to permit the reticuloendothelial system to recognize the cell as abnormal.

Most symptomatic patients have a moderate to severe anemia [hemoglobin levels 60 to 100 g/L and reticulocyte counts 10 to 30% (200 to 600 ´10_3/uL)], spherocytosis (Plate V-8), and splenomegaly.

Severe immunohemolytic anemia presents with fulminant hemolysis associated with hemoglobinemia, hemoglobinuria, and shock; this syndrome may be rapidly fatal unless aggressively treated.

The direct Coombs test is positive in 98% of patients; usually IgG is detected with or

without C3. Rarely, the cells may be agglutinated by the antibody, causing difficulty in analysis by flow cytometry.

Immune thrombocytopenia also may be present (*Evans's syndrome*), a disorder in which separate antibodies are directed against platelets andRBC. Occasionally, venous thrombosis occurs.

PATHOGENESIS IgG antibodies lyse RBCby two mechanisms: (1) immune adherence of RBC to phagocytes mediated by the antibody and by complement components that become fixed to the membrane (by far the more important mechanism of destruction), and (2) complement activation. Upon binding to Fc receptors on macrophages, the antibody-coated red cell is engulfed and destroyed. If internalization is only partial, the RBC membrane is removed, resulting in the formation of spherocytes, which are destroyed in the spleen. Complement-mediated immune adherence involves the interaction of C3b and C4b with receptors on the macrophage; while much less likely to lead to RBC lysis, this mechanism markedly increases the immune adherence due to IgG. Immune adherence, particularly that due to the IgG antibody, is also enhanced by the transit of RBC into the cords and sinuses of the spleen, which brings cells into intimate contact with phagocytic cells.

TREATMENT

Patients having a mild degree of hemolysis usually do not require therapy. In those with clinically significant hemolysis, initial therapy consists of glucocorticoids (e.g., prednisone, 1.0 mg/kg per day). A rise in hemoglobin is frequently noted within 3 or 4 days and occurs in most patients within 1 to 2 weeks. Prednisone is continued until the hemoglobin level has risen to normal values, and thereafter it is tapered rapidly to about 20 mg/d, then slowly over the course of several months. An algorithm for this tapering process is given inFig. 108-4. For chronic therapy with prednisone, alternate-day administration is preferred. More than 75% of patients achieve an initial significant and sustained reduction in hemolysis; however, in half these patients the disease recurs, either during glucocorticoid tapering or after its cessation. Glucocorticoids have two modes of action: an immediate effect due to inhibition of the clearance of IgG-coatedRBC by the mononuclear phagocyte system and a later effect due to inhibition of antibody synthesis. Splenectomy is recommended for patients who cannot tolerate or fail to respond to glucocorticoid therapy.

Patients who have been refractory to glucocorticoid therapy and to splenectomy are treated with immunosuppressive drugs such as azathioprine and cyclophosphamide. A success rate of ~50% has been reported with each. Intravenous gamma globulin may cause rapid cessation of hemolysis; however, it is not nearly as effective in this disorder as in immune thrombocytopenia.

Patients with severe anemia may require blood transfusions. Because the antibody in this disease is usually a "panagglutinin," reacting with nearly all normal donor cells, cross-matching is impossible. The goal in selecting blood for transfusion is to avoid administeringRBC with antigens to which the patient may have alloantibodies. A common procedure is to adsorb the panagglutinin present in the patient's serum with the patient's own RBC from which antibody has been previously eluted. Serum cleared

of autoantibody can then be tested for the presence of alloantibody to donor blood groups. ABO-compatible RBC matched in this fashion are administered slowly, with watchfulness for signs of an immediate-type hemolytic transfusion reaction.

PROGNOSIS In most patients, hemolysis is controlled by glucocorticoid therapy alone, by splenectomy, or by a combination. Fatalities occur among three rare subsets of patients: (1) those with overwhelming hemolysis who die from anemia; (2) those whose host defenses are impaired by glucocorticoids, splenectomy, and/or immunosuppressive agents; and (3) those with major thrombotic events coincident with active hemolysis.

When immunohemolysis develops as a complication of an underlying disorder, the prognosis is often dominated by that of the primary disease.

Immunohemolytic anemia secondary to drugs Drugs cause immunohemolytic anemia by two mechanisms of action: (1) they induce a disorder identical in almost every respect to warm-antibody immunohemolytic anemia (e.g.,a-methyldopa (an antihypertensive;Chap. 246), and (2) they become associated as haptenes with theRBCsurface and induce the formation of an antibody directed against the RBC-drug complex (e.g., penicillin, quinidine).

A positive direct Coombs test is observed in up to 10% of patients receivinga-methyldopa therapy in doses of 2 g/d or higher. A small minority of these patients develop spherocytosis and hemolysis, which may be severe. a-Methyldopa alters and makes immunogenic the protein(s) of the Rh complex; the resulting antibodies cross-react with the normal Rh protein. Thus the antibody does not react with the drug, and the indirect Coombs test is positive in almost all patients even when the drug is not added to the test. TheRBC are coated with IgG but not C3. Hemolysis decreases over the course of several weeks after cessation of drug therapy, although the direct Coombs test may remain positive for more than 1 year.

In most other cases of drug-induced hemolysis, the antibody is directed against the combination of the drug and the membrane glycoprotein to which it is attached. The hemolytic reaction in vivo is dependent on the presence of the drug and usually ceases shortly after the drug has been discontinued. Penicillin and its congeners may cause this type of reaction if the drug is given in very high doses (10 million units per day or more). The drug adheres relatively firmly to the protein of theRBCmembrane. Complement is not usually fixed, and the hemolysis in vivo is usually not severe. Since the antibody is usually IgG, spherocytosis and splenic destruction may occur. Most other drugs (such as quinine, quinidine, sulfonamides, sulfonureas, phenacetin, stibophen, and dipyrone) do not adhere as tightly to their glycoproteins, and the drug-antibody complexes are removed during the washing steps of the direct and indirect Coombs reactions. These antibodies (particularly IgM) are usually able to fix complement, and these components remain on the RBC surface; thus the direct Coombs test is positive with anti-C3 but not anti-IgG. The antibody is detected in the *indirect* Coombs test only when the drug is added to the incubation mixture. Hemolysis may be quite severe, sometimes resulting in signs of intravascular hemolysis; resolution is usually prompt after the drug is discontinued.

Immune hemolysis due to cold-reactive antibodies Antibodies that react with

polysaccharide antigens are usually IgM and react better at temperatures lower than 37°C, hence the name *cold-reactive antibodies*. Uncommonly, the antibody is IgG (the Donath-Landsteiner antibody of paroxysmal cold hemoglobinuria).

Cold agglutinins arise in two clinical settings: (1) monoclonal antibodies, the product of lymphocytic neoplasia or paraneoplasia, and (2) polyclonal antibodies in response to infection. In many elderly patients, the "neoplasm" is benign monoclonal gammopathy that does not progress, and the paraprotein remains its only manifestation. Occasionally, cold agglutinins are found in patients with nonlymphoid neoplasms.

Transient cold agglutinins occur commonly in two infections: *Mycoplasma pneumoniae* infection and infectious mononucleosis. In both, the titer of antibody is usually too low to cause clinical symptoms, but its presence is of diagnostic value; only occasionally is hemolysis present. Cold agglutinins are less frequently encountered in a number of other viral infections. Their manifestations are usually benign.

The specificity of the antibody may be of diagnostic value. Cold agglutinins reacting more strongly with adultRBC than fetal (cord) RBC are called *anti-I*; these antibodies are seen in benign lymphoproliferation (chronic cold agglutinin monoclonal gammopathy) and in *Mycoplasma* infections. Those reacting more strongly with cord RBC cells are called *anti-i*; these antibodies are seen in aggressive lymphomas and in infectious mononucleosis. Rarely, the antibody may react with other antigens that are equally expressed on adult and cord RBC. The clinical manifestations elicited by the antibody on exposure to cold are of two sorts: intravascular agglutination (acrocyanosis) and hemolysis. Acrocyanosis is the marked purpling of the extremities, ears, and nose when the blood becomes cold enough to agglutinate in the veins; it clears on warming and does not have the vasospastic characteristics of Raynaud's phenomenon (Chap. 248). Patients may also have symptoms when swallowing cold food or drinks.

The hemolysis is usually not severe and is manifested by a mild reticulocytosis, agglutination on the blood film, and agglutination during analysis of the blood by particle analysis (giving rise to a falsely high mean corpuscular volume). The degree of hemolysis depends on several variables.

1. *Antibody titer*. In general, the titer in symptomatic patients is above 1:2000 dilution of serum and may range to as high as 1:50,000. When collecting samples, great care must be taken that the serum is separated from the cells while the sample is maintained at 37°C so that the antibody will not adsorb onto the patient's own cells.

2. *Thermal amplitude of the antibody* (the highest temperature at which the antibody will react with theRBC). For most antibodies, this is 23 to 30°C. Those with a higher thermal amplitude (up to 37°C) are more hemolytic, since it is more likely that these temperatures will be encountered during RBC circulation.

3. *Environmental temperature*. Since the reaction can occur only at temperatures below body temperature, frequency and degree of exposure to cold are major determinants of the rate of hemolysis.

The hemolysis that occurs is due primarily to the hemolytic action of complement, since

there are no functional Fc receptors for the IgM antibody. Complement is readily fixed; a single molecule of IgM is enough to effect binding of C1 and initiate the cascade. However, the normal human RBC is remarkably resistant to the hemolytic action of complement because of several defense mechanisms. Therefore, severe hemolysis with hemoglobinuria occurs only with massive activation of the antibody, such as by sudden cooling. The activation of complement is always marked by the accumulation of a degradation product of C3, C3dg, on the surface; this product is what is detected with appropriate antisera in the direct Coombs test in all patients with significant cold agglutinin disease. The cutaneous manifestations and hemolysis are best treated by maintaining the patient in a warm environment.

Splenectomy is usually not of value in this disorder. Glucocorticoids are of limited value, although patients with the panthermal variety of cold agglutinin disease may respond. Chlorambucil and cyclophosphamide are commonly used to treat patients who have hemolysis associated with monoclonal gammopathy, but their efficacy is usually marginal. Successful treatment of the malignant neoplasm responsible for the cold agglutinin often reduces the titer of antibody and the severity of the hemolysis.

Chronic cold agglutinin disease tends to be unremitting. The overall prognosis is dominated by the underlying lymphoproliferative disease, if present. In those patients in whom cold agglutinin disease appears to arise spontaneously, malignant lymphoma may develop after several years.

Paroxysmal cold hemoglobinuria (PCH) Now a rare disorder, PCH was more frequent when tertiary syphilis was prevalent; now, most cases are either secondary to a viral infection or are autoimmune. PCH results from the formation of the Donath-Landsteiner antibody, an IgG antibody that is directed against the P antigen (Chap. 114) and that can induce complement-mediated lysis. Attacks are precipitated by exposure to cold and are associated with hemoglobinemia and hemoglobinuria; chills and fever; back, leg, and abdominal pain; headache; and malaise. Recovery from the acute episode is prompt, and between episodes patients are usually asymptomatic. When this syndrome accompanies acute viral infections (e.g., measles and mumps in children), it is self-limited but may be severe. Although the direct Coombs test may show complement to be present (seldom IgG), this test may be negative. The diagnosis is made by demonstrating cold-reacting IgG antibodies either by lytic tests (when the titer is very high) or by special antiglobulin tests. When PCH is secondary to syphilis, it responds to therapy for syphilis. Chronic autoimmune PCH may respond to prednisone or cytotoxic therapy (azathioprine or cyclophosphamide) but does not respond to splenectomy. The natural history of this disease often extends over many years.

Hemolysis due to Trauma in the Circulation RBC may be fragmented by mechanical trauma as they circulate; this circumstance leads to intravascular hemolysis and in most cases to RBC fragments called *schistocytes*. Schistocytes are identified by the sharp points that result from the faulty resealing of the fractured membrane (Plate V-28). Mechanical trauma leading to hemolysis occurs in three clinical settings: (1) when RBC flow through small vessels over the surface of bony prominences and are subject to external impact during various physical activities, (2) when RBC flow across a pressure gradient created by an abnormal heart valve or valve prosthesis (macrovascular), and (3) when the deposition of fibrin or small platelet thrombi in the microvasculature

exposes RBC to a physical impediment that fragments them (microvascular) (Table 108-8).

External impact Hemoglobinemia and hemoglobinuria have been observed in a small proportion of individuals who have undergone a prolonged march or a prolonged run, most typically on a hard surface and while wearing thin-soled shoes. The role of direct external trauma in this process has been demonstrated by the fact that hemolysis can be prevented by the insertion of a soft inner sole in the runner's shoes. Similar types of hemolysis have been described following karate and the playing of bongo drums. No abnormality ofRBC has been demonstrated, even during the acute episode. Susceptible individuals will develop hemoglobinemia and hemoglobinuria when exposed to the conditions described above. Muscle damage during some of these activities may produce myoglobinuria, but renal function is preserved. No specific therapy is required except to obtain better running shoes.

Macrovascular traumatic hemolysis Hemolysis associated with fragmentedRBC(Plate V-28) occurs in approximately 10% of patients with artificial aortic valve prostheses. This incidence is somewhat greater with valves having stellite rather than Silastic occluders, greater with small valves as compared with larger valves, and greater when valves are cloth-covered or when there is a paravalvular leak. Traumatic hemolysis is rare in recipients of porcine valves. Severe hemolysis may occur after repair of ostium primum or endocardial cushion defects with a prosthetic patch. Mitral valve prostheses may produce hemolysis, but since the pressure gradient across these valves is lower than across aortic prostheses, the incidence is lower. A moderately shortened RBC survival time with little or no anemia occurs in some patients with severe calcific aortic stenosis. Indeed, almost any intracardiac lesion that alters hemodynamics may lead to some shortening of RBC survival. Traumatic hemolysis has been observed in patients who have undergone aortofemoral bypass.

CLINICAL MANIFESTATIONS In severe cases, hemoglobin levels fall to 50 to 70 g/L with reticulocytosis, fragmentedRBC in the peripheral blood, depressed haptoglobin, elevated serumLDH, and hemoglobinemia and hemoglobinuria. Iron loss (as hemoglobin or hemosiderin) in the urine may lead to iron deficiency. The direct Coombs test may rarely become positive.

PATHOGENESIS A number of factors combine to cause the fragmentation ofRBC by prostheses: (1) the shear stress resulting from turbulent blood flow, particularly when blood is forced through a small aperture by high pressure (e.g., a paravalvular leak around an aortic valve); (2) direct mechanical trauma of RBC at the time of seating of the occluder of the prosthetic valve; and (3) the deposition of fibrin across disrupted attachment points.

TREATMENT

Iron deficiency should be corrected by the administration of oral iron. The elevated hemoglobin that results may permit a decrease in the cardiac output and a slowing of the hemolytic rate. Limitation in physical activity also lessens the hemolytic rate. When these measures fail, any paravalvular leak must be repaired or the prosthetic valve replaced.

Microvascular traumatic hemolysis If fibrin or platelet microthrombi are deposited in arteriolar sites,RBC may be trapped on the meshwork and fragmented by high shear forces.

ABNORMALITIES OF THE VESSEL WALL Disorders such as malignant hypertension, eclampsia, renal allograft rejection, disseminated cancer, hemangiomas, or disseminated intravascular coagulation (DIC) may cause traumatic hemolysis. The degree of hemolysis induced by this family of disorders is usually quite mild, but a large number of fragments may be seen in the peripheral blood. In some patients, thrombocytopenia may be severe. Therapy is best directed at the primary disease. Thus, reversal of renal graft rejection, treatment of malignant hypertension and eclampsia, control of cancer, and the like, lead to a cessation of hemolysis. The relative importance of the primary vascular abnormality versus fibrin deposition is unclear.

Thrombotic thrombocytopenia purpura (TTP) This disorder is characterized by arteriolar lesions in various organs that contain platelet thrombi and produce thrombocytopenia and hemolytic anemia due to fragmentation ofRBC. Tissue hypoxia resulting from vessel occlusion may cause organ dysfunction, most frequently manifest in the nervous system or the kidney. The disease affects individuals of all ages, but primarily young adults and more often women.

CLINICAL MANIFESTATIONS The classic pentad ofTTPincludes hemolytic anemia with fragmentation of erythrocytes and signs of intravascular hemolysis, thrombocytopenia, diffuse and nonfocal neurologic findings, decreased renal function, and fever. These signs and symptoms occur variably, depending on the number and sites of the arteriolar lesions. The anemia may be very mild to very severe, and the thrombocytopenia often parallels it. The neurologic and renal symptoms are usually seen only when the platelet count is markedly diminished (<20 to 30´10^3/uL). Fever is not reliably present. TTP may be acute in onset, but its course spans days to weeks in most patients and occasionally continues for months. Proteinuria and a moderate elevation of blood urea nitrogen may be found on initial presentation; the latter continues to rise while urine output falls if the patient develops renal failure. Neurologic symptoms develop in>90% of patients whose disease terminates in death. Initially, changes in mental status such as confusion, delirium, or altered states of consciousness may occur. Focal findings include seizures, hemiparesis, aphasia, and visual field defects. These neurologic symptoms may fluctuate and terminate in coma. Involvement of myocardial blood vessels may be a cause of sudden death. The severity of the disorder can be estimated from the degree of anemia and thrombocytopenia and the serumLDHlevel. Prothrombin time, partial thromboplastin time, fibrinogen concentration, and the level of fibrin split products are usually normal or only mildly abnormal. If the coagulation tests indicate a major consumption of clotting factors, the diagnosis of TTP is doubtful. A positive antinuclear antibody (ANA) determination is obtained in approximately 20% of patients.

PATHOGENESIS The manifestations ofTTP can be explained by *localized* platelet thrombi. The agglutination of platelets is mediated by unusually large multimers of von Willebrand factor. Patients with TTP have acquired an antibody that inhibits a protease that normally cleaves von Willebrand factor. Arterioles are filled with hyalin material, presumably fibrin and platelets, and similar material may be seen beneath the

endothelium of otherwise uninvolved vessels. Immunofluorescence studies have shown the presence of immunoglobulin and complement in arterioles. Microaneurysms of arterioles are often present. An association with pregnancy, AIDS, systemic lupus erythematosus (SLE), scleroderma, and Sjogren's syndrome suggests an immunologic origin.

DIAGNOSIS The combination of hemolytic anemia with fragmentedRBC, thrombocytopenia, normal coagulation tests, fever, neurologic disorders, and renal dysfunction is virtually pathognomonic ofTTP. Although they are not usually required for diagnosis, biopsies of skin and muscle, gingiva, lymph node, or bone marrow may show the typical arteriolar abnormalities. TTP must be distinguished from idiopathic thrombocytopenic purpura or Evans's syndrome (the former plus immunohemolytic anemia) by the finding of fragmented but not spherocytic RBC in the peripheral blood and a negative direct Coombs test.

TREATMENT

Plasma exhange permits >90% of patients to survive if therapy is promptly instituted. Many patients require daily or even twice daily plasmapheresis with plasma replacement. If a response is obtained (as indicated by increasing platelet count and decreasing plasmaLDH and fragmentedRBC), plasmapheresis may be done less frequently but sometimes must be continued for several weeks to months. Most patients also receive high doses of glucocorticoids and some receive platelet-active agents (dipyridamole, sulfinpyrazone, dextran, aspirin), but their efficacy is not proven. Vincristine, cyclophosphamide, or splenectomy has been used to treat patients who do not respond to plasma exchange. Even coma is not a contraindication to therapy, since full neurologic recovery is the rule in patients responding to therapy. Relapses have been noted in ~10% of patients but are usually responsive to retreatment. Platelet transfusions should not be given because they can precipitate thrombotic events.

Hemolytic-uremic syndrome This disorder is similar toTTP and is characterized by the same arteriolar lesions, which may be confined to the kidney, and by similar laboratory findings. It is usually encountered in young children. Often the patient has a prodrome of a gastroenteritic bloody diarrhea caused by *Escherichia coli* 0157:H7, and the lesions are thought to be due to the elaboration of Shiga-like verotoxins that damage renal vascular endothelial cells. This disorder has been associated with eating undercooked meat. Very rarely, the disorder appears to be familial. Patients present with acute hemolytic anemia, thrombocytopenic purpura, and acute oliguric renal failure. Most patients have either hemoglobinuria or anuria. Unlike TTP, neurologic manifestations are uncommon. The peripheral blood and coagulation tests are usually indistinguishable from those of TTP.

Patients are treated with plasmapheresis, dialysis, and transfusions. The efficacy of glucocorticoids, dextran, and heparin is uncertain. The mortality rate in children ranges from 5 to 20% but is considerably higher in adults. A disorder resembling the hemolytic-uremic syndrome has been described in adults treated with the antineoplastic drug mitomycin C, usually in combination with other drugs. It may also occur in patients receiving high-dose chemotherapy with autologous stem cell transplantation.

Disseminated intravascular coagulation Inappropriate activation of the clotting system with deposition of fibrin in small vessels may lead toRBCfragmentation in the microvasculature. RBC fragmentation occurs in about one-fourth of patients withDIC(Chap. 117). The degree of hemolysis is much less in DIC than in eitherTTP or the hemolytic-uremic syndrome, and anemia with reticulocytosis is rare.

Environmental Alteration of the Red Cell Membrane by "Toxic" Effects A variety of infections may be associated with severe hemolysis. The microbes that cause bartonellosis (Chap. 163), as well as malaria and babesiosis (Chap. 214) directly parasitizeRBC. *Clostridium welchii* (Chap. 145) produces a phospholipase that can cleave the phosphoryl bond of lecithin, thereby lysing human RBC. A mild, transient hemolysis frequently accompanies bacteremia with diverse organisms such as pneumococci, staphylococci, and *E. coli*.

Hemolysis may result from the direct action of snake and spider venoms on theRBC. Although cobra venom is directly lytic in vitro, the clinical disease induced by the bite of the cobra is one of moderate hemolysis associated with spherocytosis. Spider bites, particularly the bite of the brown recluse spider, induce acute intravascular hemolysis associated with spherocytosis and fragments of complement components on the RBC. The hemolysis continues for several days up to 1 week.

Copper has a direct hemolytic effect on RBC. Hemolysis has been observed after exposure of individuals to copper salts (such as during hemodialysis). Transient episodes of hemolysis occur in patients with Wilson's disease (Chap. 348).

TheRBCmembrane is unstable at temperatures above 49°C due to denaturation of the cytoskeletal protein spectrin. The RBC undergoes a process of budding, cleavage, and resealing above this temperature. Patients with extensive burns have prominent spherocytosis, hemoglobinemia, and sometimes hemoglobinuria.

Spur Cell Anemia Hemolytic anemia with bizarre-shapedRBCoccurs in about 5% of patients with severe hepatocellular disease, particularly advanced Laennec's cirrhosis.

Clinical manifestations Anemia is more severe than is observed in otherwise uncomplicated cirrhosis. Hematocrit levels range between 15 and 25%. Splenomegaly is always present and is greater than in patients who have cirrhosis without spur cell anemia. Jaundice may be severe because of the hemolysis and liver dysfunction, and hepatic encephalopathy is common. TheRBC are irregularly shaped with multiple spicules, and a small number of bizarre-shaped fragments are commonly seen on peripheral blood smears (Plate V-27). Reticulocytosis and other signs of hemolysis are present. The tests of liver function are typical of patients with severe cirrhosis.

RBChalf-life is decreased to as short as 6 days (normal being 26 to 32 days); RBC destruction is localized to the spleen. Normal transfused RBC acquire the defect and have a survival time similar to that of the patient's own RBC.

Pathogenesis The surface membrane of a spur cell contains 50 to 70% excess cholesterol, but its total phospholipid content is normal. By contrast, the target-shapedRBC is more common in liver disease and has an excess of both

cholesterol and phospholipid. The selective cholesterol excess in the spur cell is due to abnormal low-density lipoprotein with an increased mole ratio of free (unesterified) cholesterol to phospholipid. Cholesterol out of proportion to phospholipid decreases the fluidity of the spur cell membrane, and cell deformability is decreased. These rigid, cholesterol-laden RBC cannot pass through the filtering system of the spleen, further impeded by congestive splenomegaly in cirrhosis.

Diagnosis Patients with spur cell anemia have severe hemolysis and characteristicRBCmorphology. Increasing anemia in a patient with chronic cirrhosis most commonly results from blood loss, folic acid deficiency, or iron deficiency.

RBCof similar morphology are seen in patients with abetalipoproteinemia. However, hemolysis is minimal.

Spur cells or acanthocytes have irregular spikes (irregular in length of projections and their spacing) and must be distinguished from regularly spaced, crenatedRBC(echinocytes). Echinocytes are a frequent artifact on portions of some blood smears, and they are uniformly present in some patients with uremia ("burr cells") (Plate V-3). Small, dense crenated spheres (spheroechinocytes) are sometimes seen in congenital nonspherocytic hemolytic anemia due to enzyme deficiencies in the Embden-Meyerhof pathway (see above).

TREATMENT

Transfusion therapy is of limited benefit. Attempts to influenceRBCcholesterol with various lipid-lowering agents have been unsuccessful. Splenectomy has been reported to prevent both the conditioning of RBC in the spleen and their premature destruction. However, splenectomy carries a high risk in patients with severe liver disease complicated by portal hypertension and coagulation defects. It must be reserved for patients in whom hemolysis is a major problem and who are relatively good surgical risks.

Prognosis Spur cell anemia occurs during the late stages of cirrhosis, and>90% of patients succumb to their underlying liver disease within 1 year of the diagnosis of spur cell anemia.

Paroxysmal Nocturnal Hemoglobinuria (PNH) This hemolytic disorder is distinctive because it is an intracorpuscular defect acquired at the stem cell level.

Clinical manifestations The three common manifestations ofPNH are: hemolytic anemia, venous thrombosis, and deficient hematopoiesis. Anemia is highly variable with hematocrit values ranging from£20% to normal.RBC are normochromic and normocytic unless iron deficiency has occurred from chronic iron loss in the urine.

Granulocytopenia and thrombocytopenia are common and reflect deficient hematopoiesis. Clinical hemoglobinuria is intermittent in most patients and never occurs in some, but hemosiderinuria is usually present. The lack of two proteins, decay-accelerating factor (DAF, CD55) and a membrane inhibitor of reactive lysis (MIRL, CD59) (see below) make theRBC more sensitive to the lytic effect of

complement.

DAF normally disrupts the enzyme complexes from either the classical (antibody-driven) pathway or the alternative pathway that activate C3 and C5; CD59 inhibits the conversion of C9 by the membrane attack complex C5b-8 to a polymeric complex capable of penetrating the membrane.

The platelets also lack these proteins, but the life span of the platelet is normal. However, the activation of complement indirectly stimulates platelet aggregation and hypercoagulability; this probably accounts for the tendency to thrombosis seen inPNH.

Venous thrombosis is a common complication of patients of European origin, affecting ~40% at one time or another; it is less common in Asian patients. It occurs primarily in intraabdominal veins (hepatic, portal, mesenteric) and results in the Budd-Chiari syndrome, congestive splenomegaly, and abdominal pain. It may occur in cerebral venous sinuses and is a common cause of death in patients withPNH. The bone marrow may appear normocellular, but in vitro marrow progenitor assays are abnormal. In about 15 to 30% of long-term survivors of aplastic anemia, PNH cells appear in the circulation; in some patients, the manifestations of PNH become dominant. Patients with PNH may have aplastic periods lasting from weeks to years. PNH may be seen in association with other stem cell disorders, including myelofibrosis, and (rarely) other myelodysplastic or myeloproliferative disorders.

Pathogenesis PNH is an acquired clonal disease, arising from an inactivating somatic mutation in a single abnormal stem cell of a gene on the X-chromosome (*pig-A*) important for the biosynthesis of the glycosylphosphatidylinositol (GPI) anchor. This anchor is necessary for the attachment of a number of proteins to the external membrane surface, and its partial or complete absence results in the absence of those proteins; to date, about 20 proteins have been found to be missing on the blood cells of patients with PNH. The normal clone of stem cells does not completely disappear, and the proportion of cells that are abnormal varies among patients and over time in a single patient.

Diagnosis PNHshould be suspected in anyone with otherwise unexplained hemolytic anemia, especially with leukopenia and/or thrombocytopenia and with evidence of intravascular hemolysis (hemoglobinemia, hemoglobinuria, hemosiderinuria, elevatedLDH). Anyone recovering from aplastic anemia should be examined at intervals for the appearance of the diagnostic cells. The diagnosis is often delayed because (1) it is not considered, (2) hemoglobinuria is confused with hematuria, (3) elevation of the LDH is attributed to liver disease, and (4) the diagnostic tests (Ham's test and the sucrose lysis test) are not reliable.

For many years, the diagnosis ofPNHdepended on the demonstration of the lysis ofRBCafter complement activation either by acid (Ham or acidified serum lysis test) or by reduction in ionic strength (sucrose lysis test). These tests are inferior to the analysis of GPI-linked proteins (e.g., CD59, DAF) on RBC and granulocytes by flow cytometry.

TREATMENT

Transfusion therapy is useful inPNH not only for raising the hemoglobin level but also for suppressing the marrow production ofRBCduring episodes of sustained hemoglobinuria. Washed RBC are the preferred source to prevent exacerbation of hemolysis. Therapy with androgens sometimes results in a rise in hemoglobin level. Glucocorticoids reduce the rate of hemolysis in moderate doses (15 to 30 mg prednisone) on alternate days.

Iron deficiency is common. Iron replacement may exacerbate hemolysis because of the formation of many newRBC, which may be sensitive to complement. This occurrence may be minimized by giving prednisone (60 mg/d) or by suppressing the bone marrow with transfusions.

Acute thrombosis inPNH, particularly the Budd-Chiari syndrome and cerebral thrombosis, should be treated with thrombolytic agents. Heparin therapy should be instituted rapidly and maintained for several days before changing to coumadin therapy. Antithymoctye globulin (total dose of 150 mg/kg over 4 to 10 days) is often of use in treating marrow hypoplasia; prednisone counteracts the immune-complex disease that results from the administration of this foreign protein.

In patients with either hypoplasia or thrombosis who have an appropriate sibling donor, marrow transplantation should be considered early in the course of the disease. The usual conditioning programs are sufficient to eradicate the aberrant clone.

ANEMIA OF ACUTE BLOOD LOSS

The normal capacity to compensate for acute blood loss involves cardiovascular mechanisms, an adjustment in the oxygen affinity of hemoglobin, and an increase in erythropoiesis in the marrow. The signs and symptoms of blood loss relate to the volume of the blood loss and the time frame over which the hemorrhage occurs (Table 108-9). Losses of up to 20% of the blood volume are normally tolerated by redistribution of blood flow mediated by reflex venospasm, but the presence of fever or pain may interfere with this compensation. With larger losses, blood volume redistribution is not adequate to maintain normal blood pressure: initially, hypotension is only seen on standing, but with greater losses progressively greater problems are encountered in maintaining blood pressure in sitting or supine positions. If the blood loss is more gradual, plasma volume increases, but albumin production usually lags behind the fluid shifts. It may take 2 to 3 days for the liver to generate the albumin lost in a 1500-mL bleed.

The most rapid hematologic adjustment to acute blood loss is an increase in oxygen delivery to the tissues. This is first mediated by the Bohr effect, where the more acidic milieu of the hypoperfused hypoxic tissues shifts the hemoglobin oxygen dissociation curve to the right. Over several hours theRBCincrease their production of 2,3-bisphosphoglycerate, which also enhances the unloading of oxygen to tissues. These two mechanisms can substantially increase the capacity of RBC to deliver oxygen to the tissues.

The marrow response to hemorrhage is related to the generation of erythropoietin in the kidney in response to decreased oxygen tensions. A normal response depends on the production of erythropoietin, the presence of normal erythroid progenitors in the marrow,

and an adequate supply of iron. If these three elements are normal, reticulocytes begin to increase in number in the first 2 days based on early release of reticulocytes from the marrow. However, it takes 3 to 6 days for erythroid hyperplasia to appear and 7 to 10 days before the erythropoietic response is maximal, producing reticulocyte counts up to 20 to 30%, a reticulocyte index of³3, and a marked increase in the marrow erythoid/granulocytic ratio.

DIAGNOSIS

Usually it is clear that a patient is bleeding; however, in some cases, large volumes of blood loss can occur internally from the gastrointestinal tract (esophageal varicies, cancer in the stomach or colon), a ruptured spleen, fractures and other trauma, or other lesions that can cause massive hemorrhage into the peritoneal cavity, pleural cavity, or the retroperitoneal space. Patients who have bled sufficiently to develop hypotension generally develop anemia, which is apparent only after volume replacement. The granulocyte count may increase to³20,000 cells/uL and include immature cell types such as metamyelocytes and myelocytes. Epinephrine-induced demargination of peripheral granulocytes and release of cells from the marrow may account for this change. NucleatedRBC may appear in the circulation, and platelet counts may exceed 1´10⁶/uL. The basis for the increased platelet count is unclear. Hemorrhage in an internal cavity is accompanied by a rise in unconjugated bilirubin and a fall in serum haptoglobin.

TREATMENT

Treatment of the underlying cause of the hemorrhage is of paramount importance. If the patient is severely anemic or sufficiently hypovolemic, packedRBCshould be transfused. In less severe cases, if the patient has normal kidneys (and presumably a normal erythropoietin response to anemia), normal bone marrow function, and an adequate supply of iron, no specific therapy for the anemia is required.

(Bibliography omitted in Palm version)

Back to Table of Contents

109. APLASTIC ANEMIA, MYELODYSPLASIA, AND RELATED BONE MARROW FAILURE SYNDROMES - *Neal S. Young*

The hypoproliferative anemias associated with marrow damage include aplastic anemia, myelodysplasia (MDS), pure red cell aplasia (PRCA), and myelopthisis. Anemia in these disorders, which is normochromic, normocytic, or macrocytic and characterized by low reticulocyte count, is not a solitary or even the major finding in these diseases, which are better described as marrow failure states. In bone marrow failure, pancytopenia -- anemia, leukopenia, and thrombocytopenia (sometimes in various combinations) -- results from deficient hematopoiesis, as distinguished from blood count depression due to peripheral destruction of red cells (hemolytic anemias), platelets (idiopathic thrombocytopenic purpura or due to splenomegaly), and granulocytes (as in the immune leukopenias).

Hematopoietic failure syndromes are classified by dominant morphologic features of the bone marrow (Table 109-1). While practical distinction among these syndromes is clear in stereotypical cases, they can, of course, occur secondary to other diseases, and are so closely related that the differential diagnosis may be arbitrary, patients may seem to suffer from two or three related diseases simultaneously, or one diagnosis may appear to evolve into another. Finally, there is an important pathophysiologic relationship among these syndromes in their sharing of immune-mediated mechanisms of marrow destruction and some element of genomic instability resulting in a higher rate of malignant transformation.

APLASTIC ANEMIA

DEFINITION

Aplastic anemia is pancytopenia with bone marrow hypocellularity. Acquired aplastic anemia is distinguished from iatrogenic marrow aplasia, the common occurrence of marrow hypocellularity after intensive cytotoxic chemotherapy for cancer. Aplastic anemia can also be constitutional: the genetic disease Fanconi's anemia, while frequently associated with typical physical anomalies and the development of pancytopenia early in life, can also present as marrow failure in normal-appearing adults. Acquired aplastic anemia is often stereotypical in its manifestations, with the abrupt onset of low blood counts in a previously well young adult; seronegative hepatitis or a course of an incriminated medical drug may precede the onset. The diagnosis in these instances is uncomplicated. Sometimes blood count depression is moderate or incomplete, resulting in anemia, leukopenia, and thrombocytopenia in some combination. Aplastic anemia is related to both paroxysmal nocturnal hemoglobinuria (PNH;Chap. 108) and toMDS, and in some cases a clear distinction among these disorders may not be possible.

EPIDEMIOLOGY

The incidence of acquired aplastic anemia in Europe and Israel is 2 cases per million persons annually. In Thailand and China, rates of 5 to 7 per million have been established. In general, men and women are affected with equal frequency, but there is a biphasic age distribution, with the major peak among older children and young adults

and a second rise in the elderly.

ETIOLOGY

The origins of aplastic anemia have been inferred from several recurring clinical associations (Table 109-2); unfortunately, these relationships are neither a reliable guide in an individual patient nor necessarily etiologic. In addition, while most cases of aplastic anemia are idiopathic, little other than history separates these cases from those with a presumed etiology such as a drug exposure.

Radiation Marrow aplasia is a major acute sequela of radiation. Radiation damages DNA; tissues dependent on active mitosis are particularly susceptible. Nuclear accidents can involve not only power plant workers but also employees of hospitals, laboratories, and industry (food sterilization, metal radiography, etc.), as well as innocents exposed to stolen, misplaced, or misused sources. While the radiation dose can be approximated from the rate and degree of decline in blood counts, dosimetry by reconstruction of the exposure can help to estimate the patient's prognosis and also to protect medical personnel from contact with radioactive tissue and excreta.MDS and leukemia, but probably not aplastic anemia, are late effects of irradiation.

Chemicals Benzene is a notorious cause of bone marrow failure. Vast quantities of epidemiologic, clinical, and laboratory data link benzene to aplastic anemia, acute leukemia, and blood and marrow abnormalities. The occurrence of leukemia is roughly correlated with cumulative exposure, but susceptibility must also be important, as only a minority of even heavily exposed workers develop benzene myelotoxicity. The employment history is important, especially in industries where benzene is used for a secondary purpose, usually as a solvent. Benzene-related blood diseases have declined with regulation of industrial exposure. Although benzene is no longer generally available as a household solvent, exposure to its metabolites occurs in the normal diet and in the use of lead-free gasoline. The association between marrow failure and other chemicals that contain a benzene ring is much less well substantiated; these chemicals may have been contaminated with benzene in manufacture, or petroleum distillates may have been used to dissolve the product.

Drugs (SeeTable 109-3) Many chemotherapeutic drugs have marrow suppression as a major toxicity; effects are dose-dependent and will occur in all recipients. In contrast, idiosyncratic reactions to a large and diverse group of drugs may lead to aplastic anemia without a clear dose-response relationship. These associations rest largely on accumulated case reports, but a massive international study in Europe in the 1980s quantitated drug relationships, especially for nonsteroidal analgesics, sulfonamides, thyrostatic drugs, some psychotropics, penicillamine, allopurinol, and gold. Not all associations necessarily reflect causation: a drug may have been used to treat the first symptoms of bone marrow failure (antibiotics for fever or the preceding viral illness) or provoked the first symptom of a preexisting disease (petechiae by nonsteroidal anti-inflammatory agents administered to the thrombocytopenic patient). In the context of total drug use, idiosyncratic reactions, while individually devastating, are exceedingly rare events. Chloramphenicol, the most infamous culprit, reportedly produced aplasia in only about 1/60,000 therapy courses, and even this number is almost certainly an overestimate (risks are almost invariably exaggerated when based on collections of

cases; although the introduction of chloramphenicol was perceived to have created an epidemic of aplastic anemia, its diminished use was not followed by a changed frequency of marrow failure). Risk estimates are usually lower when determined in population-based studies; furthermore, the low absolute risk is also made more obvious: even a 10- or 20-fold increase in risk translates, in a rare disease, to but a handful of drug-induced aplastic anemia cases among hundreds of thousands of exposed patients.

Infections Hepatitis is the most common preceding infection, and posthepatitis marrow failure accounts for about 5% of etiologic associations in most series. Patients are usually young men who have recovered from a mild bout of liver inflammation 1 to 2 months earlier; the subsequent pancytopenia is very severe. The hepatitis is almost invariably seronegative (non-A, non-B, non-C, non-G) and presumably due to a novel, as yet undiscovered, virus. Fulminant liver failure in childhood can follow seronegative hepatitis, and marrow failure occurs at a high rate in these patients as well. Aplastic anemia can rarely follow infectious mononucleosis, and Epstein-Barr virus has been found in the marrow of a few aplastic anemia patients, some without a suggestive preceding history. Parvovirus B19, the cause of transient aplastic crisis in hemolytic anemias and of some pure red cell aplasia (see below), does not usually cause generalized bone marrow failure. Blood count depression is frequent in the course of many viral and bacterial infections but is comparatively moderate and resolves with the infection.

Immunologic Diseases Aplasia is a major consequence and the cause of death in *transfusion-associated graft-versus-host disease*, which can occur after infusion of unirradiated blood products to an immunodeficient recipient. Aplastic anemia is strongly associated with the rare collagen vascular syndrome called *eosinophilic fasciitis*, which is characterized by painful induration of subcutaneous tissues (Chap. 313). Pancytopenia with marrow hypoplasia can also occur in systemic lupus erythematosus.

Pregnancy Aplastic anemia very rarely may occur and recur during pregnancy and resolve with delivery or with spontaneous or induced abortion.

Paroxysmal Nocturnal Hemoglobinuria An acquired mutation in the *PIG-A* gene in a hematopoietic stem cell is required for the development ofPNH, but *PIG-A* mutations probably occur commonly in normal individuals. If the PIG-A mutant stem cell proliferates, the result is a clone of progeny deficient in glycosylphosphatidylinositol-linked cell surface membrane proteins (Chap. 108). Such PNH cells are now most accurately enumerated using fluorescence-activated flow cytometry of CD55 or CD59 expression on granulocytes rather than Ham or sucrose lysis tests on red cells. Deficient cells can be detected in about a quarter of patients with aplastic anemia at the time of presentation [and PNH cells are also seen in cases of hypocellularMDS (see below)]. In addition, functional studies of bone marrow from PNH patients, even those with mainly hemolytic manifestations, show evidence of defective hematopoiesis. Patients with an initial clinical diagnosis of PNH, especially younger individuals, may later develop frank marrow aplasia and pancytopenia; patients with an initial diagnosis of aplastic anemia may suffer from hemolytic PNH years after recovery of blood counts. One explanation for the aplastic anemia/PNH syndrome is selection of the deficient clones, perhaps because they are favored for proliferation in the peculiar environment of immune-mediated marrow destruction.

Congenital Disorders Fanconi's anemia, an autosomal recessive disorder, manifests as progressive pancytopenia, increased chromosome fragility, congenital developmental anomalies, and an increased risk of malignancy. Patients with Fanconi's anemia typically have short stature; cafe au lait spots; and anomalies involving the thumb, radius, and genitourinary tract. At least seven different genetic defects have been defined by complementation analysis. The most common, type A Fanconi's anemia, is due to a mutation in *FANCA*. The function of the four cloned genes so far identified in Fanconi's anemia remains unknown.

Patients with Shwachman-Diamond syndrome may develop pancreatic insufficiency, malabsorption, and neutropenia and are at risk of aplastic anemia. Dyskeratosis congenita is an X-linked disorder characterized by mucous membrane leukoplasia, dystrophic nails, reticular hyperpigmentation, and the later development of aplastic anemia in about half of patients. Mutation in the *DKC1 (dyskerin)* gene has been found in some cases.

PATHOPHYSIOLOGY

Bone marrow failure results from severe damage to the hematopoietic cell compartment. In aplastic anemia, replacement of the bone marrow by fat is apparent in the morphology of the biopsy specimen (Fig. 109-1;Plate V-13) and magnetic resonance imaging of the spine; cells bearing the CD34 antigen, a marker of early hematopoietic cells, are greatly diminished; and in functional studies, committed and primitive progenitor cells are virtually absent -- in vitro assays have suggested that the stem cell pool is reduced to 1% of normal in severe disease at the time of presentation. Qualitative abnormalities, such as limited number of operating stem cell clones or shortened telomere length, may follow from the quantitative deficiency, reflecting the shrunken and stressed state of hematopoiesis. An intrinsic stem cell defect exists for constitutional aplastic anemia, as cells from patients with Fanconi's anemia exhibit chromosome damage and death on exposure to certain chemical agents, but there is no convenient mechanism for the propagation of an *acquired* genetic abnormality that would produce a hypoproliferative (as opposed to neoplastic) disease. Aplastic anemia does not appear to result from defective stroma or growth factor production.

Drug Injury Extrinsic damage to the marrow follows massive physical or chemical insults such as high doses of radiation and toxic chemicals. For the more common idiosyncratic reaction to modest doses of medical drugs, altered drug metabolism has been invoked as a likely mechanism. The metabolic pathways of many drugs and chemicals, especially if they are polar and have limited water solubility, involve enzymatic degradation to highly reactive electrophilic compounds; these intermediates are toxic because of their propensity to bind to cellular macromolecules. For example, derivative hydroquinones and quinolones are responsible for benzene-induced tissue injury. Excessive generation of toxic intermediates or failure to detoxify the intermediates may be genetically determined and apparent only on specific drug challenge; the complexity and specificity of the pathways imply multiple susceptible loci and would provide an explanation for the rarity of idiosyncratic drug reactions.

Immune-Mediated Injury The recovery of marrow function in some patients prepared

for bone marrow transplantation with antilymphocyte globulin (ALG) first suggested that aplastic anemia might be immune-mediated. Consistent with this hypothesis was the frequent failure of simple bone marrow transplantation from a syngeneic twin, without conditioning cytotoxic chemotherapy, which also argued both *against* simple stem cell absence as the cause and *for* the presence of a host factor producing marrow failure. Laboratory data support an important role for the immune system in aplastic anemia. Blood and bone marrow cells of patients can suppress normal hematopoietic progenitor cell growth, and removal of T cells from aplastic anemia bone marrow improves colony formation in vitro. Increased numbers of activated cytotoxic T cells are observed in aplastic anemia patients and usually decline with successful immunosuppressive therapy; cytokine measurements suggest a predominant T_H1 immune response (interferon g, interleukin 2, and tumor necrosis factor). Interferon and tumor necrosis factor induce Fas expression on CD34 cells, leading to apoptotic cell death; localization of activated T cells to bone marrow and local production of their soluble factors are probably important in stem cell destruction.

Early immune system events in aplastic anemia are not well understood. Many different exogeneous antigens appear capable of initiating a pathologic immune response, but at least some of the active T cells recognize true self-antigens. The rarity of occurrence of aplastic anemia despite common exposures (medical drugs, hepatitis virus) suggests that genetically determined features of the immune response can convert a normal physiologic response into a sustained abnormal autoimmune process.

CLINICAL FEATURES

History Aplastic anemia can appear with seeming abruptness or have a more insidious onset. Bleeding is the most common early symptom; a complaint of days to weeks of easy bruising, oozing from the gums, nose bleeds, heavy menstrual flow, and sometimes petechiae will have been noticed. With thrombocytopenia, massive hemorrhage is unusual, but small amounts of bleeding in the central nervous system can result in catastrophic intracranial or retinal hemorrhage. Symptoms of anemia are also frequent, including lassitude, weakness, shortness of breath, and a pounding sensation in the ears. Infection is an unusual first symptom in aplastic anemia (unlike in agranulocytosis, where pharyngitis, anorectal infection, or frank sepsis occur early). A striking feature of aplastic anemia is the restriction of symptoms to the hematologic system, and patients often feel and look remarkably well despite drastically reduced blood counts. Systemic complaints and weight loss should point to other etiologies of pancytopenia. History of drug use, chemical exposure, and preceding viral illnesses must often be elicited with repeated questioning.

Physical Examination Petechiae and ecchymoses are often present, and retinal hemorrhages may be present. Pelvic and rectal examinations should be performed with great gentleness to avoid trauma; these will often show bleeding from the cervical os and blood in the stool. Pallor of the skin and mucous membranes is common except in the most acute cases or those already transfused. Infection on presentation is unusual but may be present if the patient has been symptomatic for a few weeks. Lymphadenopathy and splenomegaly are highly atypical of aplastic anemia. Cafe au lait spots and short stature suggest Fanconi's anemia; peculiar nails, dyskeratosis congenita.

LABORATORY STUDIES

Blood The smear shows large erythrocytes and a paucity of platelets and granulocytes. Mean corpuscular volume (MCV) is commonly increased. Reticulocytes are absent or few, and lymphocyte numbers may be normal or reduced. The presence of immature myeloid forms suggests leukemia orMDS; nucleated red blood cells suggest marrow fibrosis or tumor invasion; abnormal platelets suggest either peripheral destruction or MDS.

Bone Marrow The bone marrow is usually readily aspirated but appears dilute on smear, and the fatty biopsy specimen may be grossly pale on withdrawal; a "dry tap" suggests fibrosis or myelophthisis. In severe aplasia the smear of the aspirated specimen shows only red cells, residual lymphocytes, and stromal cells; the biopsy, which should be>1 cm in length, is superior for determination of cellularity and shows mainly fat under the microscope, with hematopoietic cells occupying, by definition, <25% of the marrow space. In the most serious cases the biopsy is virtually 100% fat. The correlation between marrow cellularity and disease severity is imperfect. Some patients with moderate disease by blood counts will have empty iliac crest biopsies, while "hot spots" of hematopoiesis may be seen in severe cases. If an iliac crest specimen is inadequate, cells should also be obtained by aspiration from the sternum. Residual hematopoietic cells should have normal morphology, except for mildly megaloblastic erythropoiesis; megakaryocytes are invariably greatly reduced and usually absent. Areas adjacent to the spicule should be searched for myeloblasts. Granulomas (in cellular specimens) may indicate an infectious etiology of the marrow failure.

Ancillary Studies Chromosome breakage studies of peripheral blood using diepoxybutane (DEB) or mitomycin C should be performed on children and younger adults to exclude Fanconi's anemia. Chromosome studies of bone marrow cells are often revealing inMDS and should be negative in typical aplastic anemia. Flow cytometric assays have replaced the Ham test for the diagnosis ofPNH. Serologic studies may show evidence of viral infection, especially Epstein-Barr virus and HIV. Posthepatitis aplastic anemia is typically seronegative. The spleen size should be determined by scanning if the physical examination of the abdomen is unsatisfactory. Magnetic resonance imaging may be helpful to assess the fat content on a few vertebrae in order to distinguish aplasia from MDS.

DIAGNOSIS

The diagnosis of aplastic anemia is usually straightforward, based on the combination of pancytopenia with a fatty, empty bone marrow. Aplastic anemia is a disease of the young and should be a leading diagnosis in the pancytopenic adolescent or young adult. When pancytopenia is secondary, the primary diagnosis is usually obvious from either history or physical examination: the massive spleen of alcoholic cirrhosis, the history or metastatic cancer or systemic lupus erythematosus, or obvious miliary tuberculosis on chest radiograph (Table 109-1).

Diagnostic problems can occur with atypical presentations and among related

hematologic diseases. While pancytopenia is most common, some patients with bone marrow hypocellularity have depression of only one or two of three blood lines, sometimes showing later progression to more recognizable aplastic anemia. The bone marrow in constitutional or Fanconi's anemia is indistinguishable morphologically from the aspirate in acquired disease. The diagnosis can be suggested by family history, abnormal blood counts since childhood, or the presence of associated anomalies of the skeletal and urogenital systems. Patients with Fanconi's anemia may have no peculiar physical findings and can present with aplastic anemia as adults, in the third and fourth decades and, rarely, even later. Aplastic anemia may be difficult to distinguish from the hypocellular variety ofMDS: MDS is favored by finding morphologic abnormalities, particularly of megakaryocytes and myeloid precursor cells, and typical cytogenetic abnormalities (see below).

PROGNOSIS

The natural history of severe aplastic anemia is rapid deterioration and death. Provision first of red blood cell and later platelet transfusions and effective antibiotics were of some benefit, but few patients showed spontaneous recovery. The major prognostic determinant is the blood count; severe disease is defined by the presence of two of three parameters: absolute neutrophil count<500/uL, platelet count<20,000/uL, and corrected reticulocyte count<1% (or absolute reticulocyte count<50,000/uL). Survival of patients who fulfill these criteria is about 20% at 1 year after diagnosis; patients with very severe disease, defined by an absolute neutrophil count<200/uL, fare even more poorly. Treatment has markedly improved survival in this disease.

TREATMENT

Treatment includes therapies that reverse the underlying marrow failure and supportive care of the pancytopenic patient. Severe acquired aplastic anemia can be cured by replacement of the absent hematopoietic cells (and the immune system) by stem cell transplant, or ameliorated by suppression of the immune system to allow recovery of the patient's residual bone marrow function. Hematopoietic growth factors have limited usefulness and glucocorticoids are of no value. Suspect exposures to drugs or chemicals should be discontinued; however, spontaneous recovery of severe blood count depression is rare, and a waiting period before beginning treatment may not be advisable unless the blood counts are only modestly depressed.

Bone Marrow Transplantation This is the best therapy for the young patient with a fully histocompatible sibling donor (Chap. 115). HLA typing should be ordered as soon as the diagnosis of aplastic anemia is established in a child or younger adult. In transplant candidates, transfusion of blood from family members should be avoided so as to prevent sensitization to histocompatability antigens; while transfusions in general should be minimized, limited numbers of blood products probably do not seriously affect outcome.

For allogeneic transplant from fully matched siblings, long-term survival rates for children are about 80%. Transplant morbidity and mortality are increased among adults, due mainly to the increased risk of chronic graft-versus-host disease and serious infections. Graft rejection was historically a major determinant of outcome in bone

marrow transplant for aplastic anemia; high rates of primary or secondary graft failure may be related to the pathophysiology of marrow failure as well as to alloimmunization from transfusions.

Most patients do not have a suitable sibling donor. Occasionally, a full phenotypic match can be found within the family and serve as well. Far more available are other alternative donors, either unrelated but histocompatible volunteers, or closely but not perfectly matched family members. Survival using alternative donors is about half that of conventional sibling transplants. These patients will be at risk for late complications, especially a higher rate of cancer, if radiation is used as a component of conditioning. The majority of adults who undergo alternative donor transplants succumb to transplant-related complications.

Immunosuppression Used alone,ALG or antithymocyte globulin (ATG) induces hematologic recovery (independence from transfusion and a leukocyte count adequate to prevent infection) in about 50% of patients. The addition of cyclosporine to either ALG or ATG has further increased response rates to about 70 to 80% and especially improved outcomes for children and for severely neutropenic patients. Combined treatment is now standard for patients with severe disease. Hematologic response strongly correlates with survival. Improvement in granulocyte number is generally apparent within 2 months of treatment. Most recovered patients continue to have some degree of blood count depression, theMCVremains elevated, and the bone marrow cellularity returns towards normal only very slowly, if at all. Relapse (recurrent pancytopenia) is frequent, often occurring as cyclosporine is discontinued; most, but not all, patients respond to reinstitution of immunosuppression, and some responders become dependent on continued cyclosporine administration. Development ofMDS, with typical marrow morphologic or cytogenetic abnormalities, occurs in about 15% of treated patients, usually but not invariably associated with a return of pancytopenia, and some patients develop leukemia. Although the laboratory diagnosis ofPNH can generally be made at the time of presentation of aplastic anemia by flow cytometry, recovered patients showing frank hemolysis or, less commonly, thrombosis should be retested for PNH. Bone marrow examinations should be performed annually or if there is an unfavorable change in blood counts.

HorseATG(ATGAM; Upjohn) is given at 40 mg/kg per day for 4 days; rabbit ALG (Thymoglobulin; SangStat), is administered at 3.5 mg/kg per day for 5 days. For ATG, anaphylaxis is a rare but occasionally fatal complication; allergy should be tested by a prick skin test with an undiluted solution and immediate observation; desensitization is feasible. ATG binds to peripheral blood cells, and therefore, platelet and granulocyte numbers may fall further during active treatment. Serum sickness, a flulike illness with a characteristic cutaneous eruption and arthralgia, often develops about 10 days after initiating treatment. Most patients are given methylprednisolone, 1 mg/kg per day for 2 weeks, to ameliorate the immune consequences of heterologous protein infusion. Excessive or extended glucocorticoid therapy is associated with avascular joint necrosis. Cyclosporine is administered orally at an initial dose of 12 mg/kg per day in adults (15 mg/kg per day in children), with subsequent adjustment according to blood levels obtained every 2 weeks. Trough levels should be between 150 and 200 ng/mL. The most important side effects of chronic cyclosporine treatment are nephrotoxicity, hypertension, seizures, and opportunistic infections, especially *Pneumocystis carinii*

(prophylactic treatment with monthly inhaled pentamidine is recommended).

Most patients with aplastic anemia lack a suitable marrow donor and immunosuppression is the treatment of choice. Long-term survival is equivalent with transplantation and immunosuppression. However, successful transplant cures marrow failure, while patients who recover adequate blood counts after immunosuppression remain at risk of relapse and malignant evolution. Because of the excellent results in children, allogeneic transplant should always be performed in the pediatric population if a suitable sibling donor is available. Increasing age and the severity of neutropenia are the most important factors weighing in the decision between transplant and immunosuppression in adults who have a matched family donor: older patients do better withATG and cyclosporine, while transplant is preferred if granulocytopenia is profound. Some reluctant patients may be treated by immunosuppression followed by transplant for failure to recover blood counts or occurrence of late complications.

Outcomes following both transplant and immunosuppression have improved with time. High doses of cyclophosphamide, without stem cell rescue, have been reported to produce durable hematologic recovery, without relapse or evolution toMDS, but this treatment can produce sustained severe neutropenia and response is often delayed. Novel immunosuppressive drugs such as mycophenolate mofetil may further improve outcome.

Other Therapies The effectiveness of androgen therapy has not been verified in controlled trials, but occasional patients will respond or even demonstrate blood count dependence on continued therapy. For patients with moderate disease or those with severe pancytopenia who have failed immunosuppression, a 3- to 4-month trial is appropriate. Hematopoietic growth factors, granulocyte colony stimulating factor (G-CSF), granulocyte-macrophage CSF (GM-CSF), and interleukin 3, are not recommended as initial therapy for severe aplastic anemia, and even their role as adjuncts to immunosuppression is not well defined. Some patients may respond to chronic administration of growth factors in combination after failing immunosuppression. Splenectomy may occasionally increase blood counts in relapsed or refractory cases.

Supportive Care Meticulous medical attention is required so that the patient may survive to benefit from definitive therapy or, having failed treatment, to maintain a reasonable existence in the face of pancytopenia. First and most important, infection in the presence of severe neutropenia must be aggressively treated by prompt institution of parenteral, broad-spectrum antibiotics, usually ceftazadime or a combination of an aminoglycoside, cephalosporin, and semisynthetic penicillin. Therapy is empirical and must not await results of culture, although specific foci of infection such as oropharyngeal or anorectal abscesses, pneumonia, sinusitis, and typhlitis (necrotizing colitis) should be sought on physical examination and with radiographic studies. When indwelling plastic catheters become contaminated, vancomycin should be added. Persistent or recrudescent fever implies fungal disease: *Candida* or *Aspergillus* are common, especially after several courses of antibacterial antibiotics, and a progressive course may be averted by timely initiation of amphotericin. Granulocyte transfusions using G-CSF-mobilized peripheral blood have been effective in the treatment of overwhelming infections in a few patients. Hand washing, the single most effective method of preventing the spread of infection, remains a neglected practice.

Nonabsorbed antibiotics for gut decontamination are poorly tolerated and not of proven value. Total reverse isolation is not clearly beneficial in reducing mortality from infections.

Both platelet and erythrocyte numbers can be maintained by transfusion. Alloimmunization limits the usefulness of platelet transfusions and can be avoided or minimized by several strategies, including use of single donors to reduce exposure and physical or chemical methods to diminish leukocytes in the product; HLA-matched platelets are often effective in patients refractory to random donor products. Inhibitors of fibrinolysis such as aminocaproic acid have not been shown to relieve mucosal oozing; the use of low-dose glucocorticoids to induce "vascular stability" is unproven. Whether platelet transfusions are better used prophylactically or only as needed remains unclear. Any rational regimen of prophylaxis requires transfusions once or twice weekly in order to maintain the platelet count>10,000/uL (oozing from the gut, and presumably also from other vascular beds, increases precipitously at counts<5000/uL). Menstruation should be suppressed either by oral estrogens or nasal follicle-stimulating hormone/luteinizing hormone (FSH/LH) antagonists. Aspirin and other nonsteroidal anti-inflammatory agents inhibit platelet function and must be avoided.

Red blood cells should be transfused to maintain a normal level of activity, usually at a hemoglobin value of 70 g/L (90 g/L if there is underlying cardiac or pulmonary disease); a regimen of 2 units every 2 weeks will replace normal losses in a patient without a functioning bone marrow. In chronic anemia, the iron chelator deferoxamine should be added at the time of the fiftieth transfusion in order to avoid secondary hemochromatosis.

PURE RED CELL APLASIA

More restricted forms of marrow failure occur, in which only a single circulating cell type is affected and the aregenerative marrow shows corresponding absence or decreased numbers of specific precursor cells: aregenerative anemia as inPRCA(see below), thrombocytopenia with amegakaryocytosis (Chap. 116), and neutropenia without marrow myeloid cells in agranulocytosis (Chap. 64). In general, and in contrast to aplastic anemia andMDS, the unaffected lineages appear quantitatively and qualitatively normal. Agranulocytosis, the most frequent of these syndromes, is usually a complication of medical drug use (with agents similar to those related to aplastic anemia), either by a mechanism of direct chemical toxicity or by immunologic mediation. Agranulocytosis has an incidence similar to aplastic anemia but is especially frequent among the elderly and in women. The syndrome should resolve with discontinuation of exposure, but significant mortality is attached to neutropenia in the older and often previously unwell patient. Both pure white cell aplasia (agranulocytosis without incriminating drug exposure) and amegakaryocytic thrombocytopenia are exceedingly rare and, like PRCA, appear to be due to destructive antibodies or lymphocytes and can respond to immunosuppressive therapies. In all the single lineage failure syndromes, progression to pancytopenia or leukemia is unusual.

DEFINITION AND DIFFERENTIAL DIAGNOSIS

PRCAis characterized by anemia, reticulocytopenia, and absent or rare erythroid

precursor cells in the bone marrow. The classification of PRCA is shown inTable 109-4. In adults, PRCA is acquired. An identical syndrome can occur constitutionally: Diamond-Blackfan anemia, or congenital PRCA, is diagnosed at birth or in early childhood and often responds to glucocorticoid treatment. Temporary red cell failure occurs in transient aplastic crisis of hemolytic anemias, due to acute parvovirus infection (Chap. 187), and in transient erythroblastopenia of childhood, which affects normal children.

CLINICAL ASSOCIATIONS AND ETIOLOGY

PRCAhas important associations with immune system diseases. A small minority of cases occur with a thymoma. More frequently, red cell aplasia can be the major manifestation of large granular lymphocytosis or may occur in chronic lymphocytic leukemia. Some patients may be hypogammaglobulinemic. As with agranulocytosis, PRCA can be due to an idiosyncratic reaction to a drug.

Like aplastic anemia,PRCAresults from diverse mechanisms. Antibodies to red blood cell precursors are frequently present in the blood, but T cell inhibition is probably the more common immune mechanism. Cytotoxic lymphocyte activity restricted by histocompatibility locus or specific for human T cell leukemia/lymphoma virus I-infected cells, as well as natural killer cell activity inhibitory of erythropoiesis, have been demonstrated in particularly well-studied individual cases.

Persistent Parvovirus B19 Infection Chronic parvovirus infection is an important, treatable cause ofPRCA. This common virus causes a benign exanthem of childhood (fifth disease) and a polyarthralgia syndrome in adults. In patients with underlying hemolysis (or any condition that increases demand for red blood cell production), parvovirus infection can cause a transient aplastic crisis and an abrupt but temporary worsening of the anemia due to failed erythropoiesis. In normal individuals, acute infection is resolved by production of neutralizing antibodies to the virus, but in the setting of congenital, acquired, or iatrogenic immunodeficiency, persistent viral infection may occur. The bone marrow shows red cell aplasia and the presence of giant pronormoblasts (Fig. 109-2), the cytopathic sign of B19 parvovirus infection and highly suggestive of the diagnosis. Viral tropism for human erythroid progenitor cells is due to its use of erythrocyte P antigen as a cellular receptor for entry. Direct cytotoxicity of virus causes anemia if demands on erythrocyte production are high; in normal individuals, the temporary cessation of red cell production is not clinically apparent, and skin and joint symptoms are mediated by immune complex deposition.

TREATMENT

History, physical examination, and routine laboratory studies may disclose an underlying disease or a suspect drug exposure. Thymoma should be sought by radiographic procedures. Tumor excision is indicated, but anemia does not necessarily improve with surgery. The diagnosis of parvovirus infection requires detection of viral DNA sequences in the blood (IgG and IgM antibodies are commonly absent). The presence of erythroid colonies has been considered predictive of response to immunosuppressive therapy in idiopathicPRCA.

Red cell aplasia is compatible with long survival with supportive care alone, a combination of erythrocyte transfusions and iron chelation. For persistent B19 parvovirus infection, almost all patients respond to intravenous immunoglobulin therapy (for example, 0.4 g/kg daily for 5 days), although relapse and retreatment may be expected, especially in patients with AIDS. The majority of patients with idiopathicPRCArespond favorably to immunosuppression. Most first receive a course of glucocorticoids, followed in the absence of a response by cyclosporine,ATG, azathioprine, or cyclophosphamide.

MYELODYSPLASIA

DEFINITION

The myelodysplastic syndromes are a heterogeneous group of hematologic disorders broadly characterized by cytopenias associated with a dysmorphic (or abnormal appearing) and usually cellular bone marrow, and consequent ineffective blood cell production (Table 109-5). The current nomenclature was developed by the French-American-British (FAB) Cooperative Group and, while increasing recognition of the syndromes, is not entirely satisfactory: chronic myelomonocytic leukemia, while associated with dysplastic morphology, behaves as a myeloproliferative disease; sideroblastic anemias likely have a distinctive etiology; and the borderline between refractory anemia with excess blasts in transformation and acute myeloid leukemia is so arbitrary as to have been abandoned in the most recent World Health Organization classification. The FAB scheme has been recently supplemented by the International Prognostic Scoring System (IPSS; Table 109-6).

EPIDEMIOLOGY

IdiopathicMDS is a disease of the elderly; the mean age at onset is 68 years. There is a slight male preponderance. MDS is a relatively common form of bone marrow failure, with incidence rates reported of 35 to >100 per million persons in the general population and 120 to >500 per million in the aged. MDS is rare in children, but monocytic leukemia can be seen. Therapy-related MDS is not age-related and may occur in as many as 15% of patients within a decade following intensive combined modality treatment for cancer. Rates of MDS have increased over time, due to the recognition of the syndrome by physicians and the aging of the population.

ETIOLOGY AND PATHOPHYSIOLOGY

The myelodysplastic syndromes have been convincingly linked to environmental exposures such as radiation and benzene; other risk factors have been reported inconsistently. SecondaryMDSoccurs as a stereotypical late toxicity of cancer treatment, usually with a combination of radiation and the radiomimetic alkylating agents such as busulfan, nitrosourea, or procarbazine (with a latent period of 5 to 7 years) or the DNA topoisomerase inhibitors (2 years). Both acquired aplastic anemia following immunosuppressive treatment and Fanconi's anemia can evolve into MDS.

MDSis a clonal hematopoietic stem cell disorder leading to impaired cell proliferation and differentiation. Cytogenetic abnormalities are found in about half of patients, and

some of the same specific lesions are also seen in frank leukemia; deletions are more frequent than translocations. Both presenting and evolving hematologic manifestations result from the accumulation of multiple genetic lesions, loss of tumor suppressor genes, activating oncogene mutations, or other harmful alterations. Cytogenetic abnormalities are not random (loss of all or part of 5, 7, and 20, trisomy of 8) and may be related to etiology (11q23 following topoisomerase II inhibitors); chronic myelomonocytic leukemia is often associated with t(5;12) that creates a chimeric *tel-PDGFb* gene. The type and number of cytogenetic abnormalities strongly correlate with the probability of leukemic transformation and survival. Mutations of N-*ras* (an oncogene), *p53* and *IRF-1* (tumor suppressor genes), *Bcl-2* (an antiapoptotic gene), and others have been reported in some patients but may occur relatively late in the sequence leading to leukemic transformation. Apoptosis of marrow cells is increased in MDS, presumably due to these acquired genetic alterations or possibly to an overlaid immune response. Sideroblastic anemia may be related to mutations in mitochondrial genes. Ineffective erythropoiesis and disordered iron metabolism are the functional consequences of the genetic alterations.

CLINICAL FEATURES

Anemia dominates the early course. Most symptomatic patients complain of the gradual onset of fatigue and weakness, dyspnea, and pallor, but at least half the patients are asymptomatic andMDS is discovered only incidentally on routine blood counts. Previous chemotherapy or radiation exposure is an important historic fact. Fever and weight loss should point to a myeloproliferative rather than myelodysplastic process. Children with Down's syndrome are susceptible to MDS, and a family history may indicate a hereditary form of sideroblastic anemia or Fanconi's anemia.

The physical examination is remarkable for signs of anemia; about 20% of patients have splenomegaly. Some unusual skin lesions, including Sweet's syndrome (febrile neutrophilic dermatosis), have been associated withMDS.

LABORATORY STUDIES

Blood Anemia is present in the majority of cases, either alone or as part of bi- or pancytopenia; isolated neutropenia or thrombocytopenia is more unusual. Macrocytosis is common, and the smear may be dimorphic with a distinctive population of large red blood cells. Platelets are also large and lack granules. In functional studies, they may show marked abnormalities, and patients may have bleeding symptoms despite seemingly adequate numbers. Neutrophils are hypogranulated; have hyposegmented, ringed, or abnormally segmented nuclei; and contain Dohle bodies and may be functionally deficient. Circulating myeloblasts usually correlate with marrow blast numbers, and their quantitation is important for classification and prognosis. The total white blood cell count is usually normal or low, except in chronic myelomonocytic leukemia. As in aplastic anemia,MDS also can be associated with a clonal population ofPNHcells.

Bone Marrow The bone marrow is usually normal or hypercellular but in 20% of cases is sufficiently hypocellular to be confused with aplasia. No single characteristic feature of marrow morphology distinguishesMDS, but the following are commonly observed:

dyserythropoietic changes (especially nuclear abnormalities) and ringed sideroblasts in the erythroid lineage; hypogranulation and hyposegmentation in granulocytic precursors, with an increase in myeloblasts; and megakaryocytes showing reduced numbers of disorganized nuclei. Prognosis strongly correlates with the proportion of marrow blasts. Cytogenetic analysis also is important. A much more sensitive method to detect infrequent chromosome aberrations is fluorescent in situ hybridization, and gene amplification by polymerase chain reaction can detect known chromosomal translocations.

DIFFERENTIAL DIAGNOSIS

Deficiencies of vitamin B_{12} or folate should be suggested by history and excluded by appropriate blood tests; vitamin B_6deficiency can be assessed by a therapeutic trial of pyridoxine if the bone marrow shows ringed sideroblasts. Marrow dysplasia can be observed in acute viral infections, drug reactions, or chemical toxicity but should be transient. More difficult (arbitrary) are the distinctions between hypocellularMDS and aplasia or between refractory anemia with excess blasts in transformation and early acute leukemia.

PROGNOSIS

The median survival varies greatly withFAB type and, according toIPSScalculations, ranges from years for patients with 5q- or sideroblastic anemia to a few months in refractory anemia with excess blasts or severe pancytopenia associated with monosomy 7. Most patients die as a result of complications of pancytopenia and not due to leukemic transformation; perhaps one-third will succumb to other diseases unrelated to theirMDS. Precipitous worsening of pancytopenia, acquisition of new chromosomal abnormalities on serial cytogenetic determination, and increase in the number of blasts are all poor prognostic indicators. The outlook in therapy-related MDS, regardless of FAB type, is very poor, and most patients will progress within a few months to refractory acute myeloid leukemia.

TREATMENT

The therapy ofMDS is generally unsatisfactory. Only stem cell transplantation offers cure: survival rates of 40% have been reported, but older patients are particularly prone to develop treatment-related mortality and morbidity. Those with better prognostic features (and a more favorable natural history) have much better outcomes than patients with more malignant subtypes. Surprisingly, results of transplant using matched unrelated donor are comparable, although most series contain younger and more highly selected cases.

MDShas been regarded as particularly refractory to cytotoxic chemotherapy regimens but is probably no more resistant to effective treatment than acute myeloid leukemia in the elderly, in whom drug toxicity is often fatal and remissions, if achieved, are brief. Low doses of cytotoxic drugs have been administered for their "differentiating" potential: responses to cytosine arabinoside did not translate into a survival advantage; etoposide and 5-azacytidine are under active study. Amifostine, an organic thiophosphonate that blocks apoptosis, can improve blood counts but has significant toxicities.

Immunosuppressive therapies, includingATG and cyclosporine, that are effective in aplastic anemia may induce sustained remissions in a high proportion of patients with refractory anemia, especially in those with hypocellular marrows or without cytogenetic abnormalities.

Hematopoietic growth factors can improve blood counts but, as in most other marrow failure states, have been most beneficial in patients with the least severe pancytopenia. G-CSF treatment alone failed to improve survival in a controlled trial. The combination of G-CSF and erythropoietin increased blood counts in one-third to one-half of patients, but survival advantage is not yet proven.

The same principles of supportive care described for aplastic anemia apply toMDS. Because many patients will be anemic for years, erythrocyte transfusion support should be accompanied by iron chelation in order to prevent secondary hemochromatosis.

MYELOPHTHISIC ANEMIAS

Fibrosis of the bone marrow (seePlate V-19), usually accompanied by a characteristic blood smear picture called *leukoerythroblastosis*, can occur as a primary hematologic disease, called *myelofibrosis* or *myeloid metaplasia* (Chap. 110), and as a secondary process, called *myelophthisis*. Myelophthisis, or secondary myelofibrosis, is reactive. Fibrosis can be a response to invading tumor cells, usually of an epithelial cancer of breast, lung, and prostate or neuroblastoma. Marrow fibrosis may occur with infection of mycobacteria (both *Mycobacterium tuberculosis* and *M. avium*) fungi, or HIV, and in sarcoidosis. Intracellular lipid deposition in Gaucher's disease and obliteration of the marrow space related to absence of osteoclast remodeling in congenital osteopetrosis also can produce fibrosis. Secondary myelofibrosis is a late consequence of radiation therapy or treatment with radiomimetic drugs. Usually, the infectious or malignant underlying processes are obvious. Marrow fibrosis can also be a feature of a variety of hematologic syndromes, especially chronic myeloid leukemia, multiple myeloma, lymphomas, myeloma, and hairy cell leukemia.

The pathophysiology has three distinct features: proliferation of fibroblasts in the marrow space (myelofibrosis); the extension of hematopoiesis into the long bones and most particularly into extramedullary sites, usually the spleen, liver, and lymph nodes (myeloid metaplasia); and ineffective erythropoiesis. The etiology of fibrosis is unknown but most likely involves dysregulated production of growth factors: platelet-derived growth factor and transforming growth factorb have been implicated. Abnormal regulation of other hematopoietins would lead to localization of blood-producing cells in nonhematopoietic tissues and uncoupling of the usually balanced processes of stem cell proliferation and differentiation. Myelofibrosis is remarkable for pancytopenia despite extraordinarily large numbers of circulating hematopoietic progenitor cells.

Anemia is dominant in secondary myelofibrosis, usually normocytic and normochromic. The diagnosis is suggested by the characteristic leukoerythroblastic smear (seePlate V-9). Erythrocyte morphology is very abnormal, with circulating nucleated red blood cells, teardrops, and shape distortions. White blood cell numbers are often elevated, sometimes mimicking a leukemoid reaction, with circulating myelocytes, promyelocytes, and myeloblasts. Platelets may be abundant and are often giant size. Inability to

aspirate the bone marrow, the characteristic "dry tap," can allow a presumptive diagnosis before the biopsy is decalcified.

The course of secondary myelofibrosis is determined by its cause, usually a metastatic tumor or an advanced hematologic malignancy. Treatable causes must be excluded, especially tuberculosis and fungus. Transfusion support can relieve symptoms.

(Bibliography omitted in Palm version)

Back to Table of Contents

110. POLYCYTHEMIA VERA AND OTHER MYELOPROLIFERATIVE DISEASES - *Jerry L. Spivak*

Polycythemia vera, idiopathic myelofibrosis, essential thrombocytosis, and chronic myeloid leukemia (CML) are commonly classified together under the rubric *the chronic myeloproliferative disorders*, because their pathophysiology involves the clonal expansion of a multipotent hematopoietic progenitor cell with the overproduction of one or more of the formed elements of the blood. These entities may transform into acute leukemia naturally or as a consequence of mutagenic treatment. However, while polycythemia vera, idiopathic myelofibrosis, essential thrombocytosis, and CML share similar phenotypic characteristics, CML is genotypically distinct from the other three disorders because it alone is associated with translocation of genetic material between the long arms of chromosomes 9 and 22, resulting in the production of the unique fusion protein, bcr-abl. Furthermore, based on its natural history, CML is more appropriately considered as a form of leukemia.**CML is discussed with the acute myeloid leukemias in Chap. 111.*

POLYCYTHEMIA VERA

Polycythemia vera is a clonal disorder involving a multipotent hematopoietic progenitor cell in which there is accumulation of phenotypically normal red cells, granulocytes, and platelets in the absence of a recognizable physiologic stimulus. Polycythemia vera, the most common of the chronic myeloproliferative disorders, occurs in about 2 per 100,000 people. It spares no adult age group. Vertical transmission has been documented, establishing a genetic basis for the disorder. A slight overall male predominance has been observed, but females predominate within the reproductive age range.

ETIOLOGY

The etiology of polycythemia vera is unknown. Although nonrandom chromosome abnormalities such as 20q-, trisomy 8 or 9 have been documented in a small percentage of untreated polycythemia vera patients, no consistent cytogenetic abnormality has been associated with the disorder and no specific genetic defect has yet been identified. Impaired posttranslational processing of the thrombopoietin receptor, Mpl, has been noted in polycythemia vera patients; the extent of the defect correlated with disease duration and splenomegaly. While this defect is specific for polycythemia vera and is not found in secondary polycythemias, its role in the pathophysiology of the disorder is still undefined. In contrast to normal erythroid progenitor cells, polycythemia vera erythroid progenitor cells can grow in vitro in the absence of erythropoietin due to hypersensitivity to insulin-like growth factor I. However, this phenotypic abnormality is not specific for polycythemia vera and has been documented in essential thrombocytosis and secondary polycythemias. Polycythemia vera erythroid progenitor cells are more resistant to apoptosis induced by erythropoietin deprivation, due to upregulation of bcl-X_L, an antiapoptotic protein. The polycythemia vera erythroid progenitors do not divide more rapidly than their normal counterparts, but they accumulate because they do not die normally. Additionally, the transformed hematopoietic progenitor cells in polycythemia vera, as in other neoplastic disorders, exhibit clonal dominance and suppress the proliferation of normal hematopoietic progenitor cells by an unknown mechanism. Consequently, the circulating formed

elements of the blood represent only progeny of the transformed clone.

CLINICAL FEATURES

Although massive splenomegaly may be the initial presenting sign in polycythemia vera, most often the disorder is first recognized by the discovery of a high hemoglobin or hematocrit, and with the exception of aquagenic pruritus, no symptoms distinguish polycythemia vera from other causes of erythrocytosis.

Uncontrolled erythrocytosis can lead to neurologic symptoms such as vertigo, tinnitus, headache, and visual disturbances. Systolic hypertension also accompanies an elevated red cell mass. In some patients, venous or arterial thrombosis may be the presenting manifestation of polycythemia vera. Intraabdominal venous thrombosis is particularly common and may be catastrophic when there is sudden compromise of the hepatic vein. Polycythemia vera should be suspected in any patient who develops the Budd-Chiari syndrome. Digital ischemia may also occur. Easy bruising, epistaxis, or gastrointestinal hemorrhage may be observed, and polycythemia vera patients are frequently hypermetabolic. Hypercuricemia with secondary gout and uric acid stones and acid-peptic disease also complicate the disorder. Because isolated erythrocytosis is a common initial presentation for polycythemia vera but no clonal marker is available for the disease, the first task of the physician is to distinguish this autonomous clonal form of erythrocytosis from the many other types of erythrocytosis, most of which are correctable (Table 110-1).

Erythropoiesis is normally regulated by the glycoprotein hormone erythropoietin. Erythropoietin, which in adults is produced primarily in the kidneys and to a small extent in the liver, promotes the proliferation of erythroid progenitor cells, maintains their survival, and facilitates their differentiation. Because erythropoietin acts as a survival factor, it is constitutively produced and, like the red cell mass, its level is constant as long as tissue oxygenation is adequate. The plasma erythropoietin level, like the red cell mass, differs among individuals but in adults is not affected by either age or gender. Erythropoietin production is regulated at the level of gene transcription. Hypoxia is the only physiologic stimulus that increases the number of cells producing erythropoietin, and thus the production and metabolism of erythropoietin are independent of its plasma level. In the absence of renal or hepatic disease, plasma erythropoietin levels reflect erythropoietin production, and therefore the assay for plasma erythropoietin is a surrogate assay for tissue hypoxia. Erythropoietin is active at the picomolar level, and its production is tightly regulated. Thus, the plasma erythropoietin level does not rise outside the normal range until the hemoglobin level falls below 105 g/L. This is not meant to imply that an increase in erythropoietin production does not occur as the hemoglobin level falls below normal, but because the normal range for plasma erythropoietin is wide (4 to 26 mU/mL), unless the patient's baseline level is known, any increase will not be recognized until the hemoglobin falls below 105 g/L. Thereafter, there is a log-linear inverse correlation between the levels of plasma erythropoietin and hemoglobin. With erythrocytosis, erythropoietin production is suppressed; this suppression reflects not only the increase in tissue oxygen transport associated with the increase in red cell number but also additional negative-feedback mechanism unrelated to oxygen transport but related to the increase in blood viscosity and an increase in red cell precursors capable of taking up erythropoietin. The summation of these

mechanisms accounts for the paradoxical observation that many patients with hypoxic erythrocytosis due to cyanotic congenital heart disease or obstructive lung disease have a "normal" plasma erythropoietin level. The plasma erythropoietin level is a useful diagnostic test in patients with isolated erythrocytosis, because an elevated level essentially excludes polycythemia vera as the cause for the erythrocytosis.

DIAGNOSIS

When confronted with an elevated hemoglobin or hematocrit level, it is important to obtain previous values to determine the duration of this laboratory abnormality. Because the hemoglobin or hematocrit level is affected by the plasma volume, and hematocrit and red cell mass are not linearly related, a red cell mass determination must also be performed to distinguish absolute erythrocytosis from relative erythrocytosis due to a reduction in plasma volume alone (also known as *stress* or *spurious erythrocytosis* or *Geisbock's syndrome*). Red cell mass determination is important because in polycythemia vera, in contrast to erythropoietin-driven erythrocytosis, the plasma volume is frequently elevated, not only masking the true extent of red cell mass expansion but often its presence. Indeed, a significant proportion of patients with polycythemia vera have a hematocrit within the normal range, particularly in patients with a substantial splenomegaly. Failure to recognize this phenomenon is undoubtedly the basis for many of the reported instances of hepatic or portal vein thrombosis in patients with a so-called undefined myeloproliferative disorder.

Red cell mass is reliably determined by isotope dilution using the patient's ^{51}Cr-tagged red cells; extrapolations made by determining directly only the plasma volume are unacceptable. Furthermore, to allow ample time for equilibration of the labeled red cells, measurements should be made over a period of ^{3}90 min.

Once the presence of absolute erythrocytosis has been established, its cause must be determined. An elevated plasma erythropoietin level suggests either an hypoxic cause for erythrocytosis or autonomous erythropoietin production, in which case assessment of pulmonary function and an abdominal computed tomography scan to evaluate renal and hepatic anatomy are appropriate. A normal erythropoietin level does not exclude an hypoxic cause for erythrocytosis. In polycythemia vera, in contrast to hypoxic erythrocytosis, the arterial oxygen saturation is normal. However, a normal oxygen saturation does not exclude a high-affinity hemoglobin as a cause for erythrocytosis, and it is here that documentation of previous hemoglobin levels and a family study become important. Because there is no clonal marker for polycythemia vera, clinical guidelines have been proposed to define the disease. A modified version is provided inTable 110-2. However, these guidelines do not establish clonality, and in some patients only with time will the underlying disorder become apparent. Diagnostic ambiguity does not preclude the initiation of therapy.

Other laboratory studies that may aid in diagnosis include the red cell count, mean corpuscular volume, and red cell distribution width (RDW). Only three situations cause microcytic erythrocytosis:b-thalassemia trait, hypoxic erythrocytosis, and polycythemia vera. However, with b-thalassemia trait the RDW is normal, whereas with hypoxic erythrocytosis and polycythemia vera, the RDW is usually elevated. A properly made blood smear from a patient with erythrocytosis will be virtually unreadable due to the

marked elevation in red cell count, but no specific morphologic abnormalities are seen in the leukocytes or platelets in polycythemia vera. However, when these are also elevated the diagnosis is assured. In many patients, the leukocyte alkaline phosphatase level is also increased, as is the uric acid level. Elevated serum vitamin B_{12} or B_{12}-binding capacity may be present. In patients with associated acid-peptic disease, occult gastrointestinal bleeding may lead to presentation with hypochromic, microcytic anemia.

A bone marrow aspirate and biopsy will provide no specific diagnostic information, and unless there is a need to establish the presence of myelofibrosis or exclude some other disorder, these procedures need not be done. Although the presence of a cytogenetic abnormality such as trisomy 8 or 9 or 20q- in the setting of an expansion of the red cell mass supports the clonal etiology, no specific cytogenetic abnormality is associated with polycythemia vera, and the absence of a cytogenetic marker does not exclude the diagnosis.

COMPLICATIONS

The major clinical complications of polycythemia vera relate directly to the increase in blood viscosity associated with elevation of the red cell mass and indirectly to the increased turnover of red cells, leukocytes, and platelets and the attendant increase in uric acid and histamine production. The latter appears to be responsible for the increase in peptic ulcer disease and for the pruritus associated with this disorder, although little formal proof for this has been obtained. A sudden massive increase in spleen size is another problem and can be associated with splenic infarction or progressive cachexia. Myelofibrosis and myeloid metaplasia can also develop with transfusion-dependent anemia, but the frequency is low in those not receiving chemotherapy or irradiation. Although acute nonlymphocytic leukemia is reported to be increased in polycythemia vera, the incidence of acute leukemia in patients not exposed to chemotherapy or radiation is low and the development of leukemia is not related to disease duration, suggesting that the treatment exposure may be a more important risk factor than the disease itself.

Erythromelalgia is a curious syndrome of unknown etiology involving primarily the lower extremities and manifested usually by erythema, warmth, and pain of the affected appendage and occasionally digital infarction. It occurs with a variable frequency in patients with a myeloproliferative disorder and is usually responsive to salicylates. Some of the central nervous system symptoms observed in patients with polycythemia vera may represent a variant of erythromelalgia.

If left uncontrolled, erythrocytosis can lead to intravascular thrombosis involving vital organs such as the liver, heart, brain, or lungs. Patients with massive splenomegaly are particularly prone to thrombotic events because the associated increase in plasma volume masks the true extent of the red cell mass elevation as measured by the hematocrit or hemoglobin level. A "normal" hematocrit or hemoglobin level in a polycythemia vera patient with massive splenomegaly should be considered as indicative of an elevated red cell mass until proven otherwise.

TREATMENT

Polycythemia vera is generally an indolent disorder whose clinical course can run many decades, and its medical management should reflect the tempo of the disorder. Maintenance of the hemoglobin level at £140 g/L in men and £120 g/L in women is mandatory to avoid the thrombotic complications. Thrombosis due to erythrocytosis is the most significant complication of this disorder. Phlebotomy serves initially to reduce hyperviscosity by bringing the red cell mass into the normal range. Periodic phlebotomies thereafter serve to maintain the red cell mass within the range of normal and to induce a state of iron deficiency, which prevents an accelerated reexpansion of the red cell mass. In most polycythemia vera patients, once an iron-deficient state is achieved, phlebotomy is usually required only at 3-month intervals. Although both phlebotomy and iron deficiency, in addition to the disease itself, tend to increase the platelet count, thrombocytosis is not correlated with thrombosis in polycythemia vera, in contrast to the strong correlation between erythrocytosis and thrombosis in this disease. The use of salicylates as a tonic against thrombosis in polycythemia vera patients is potentially harmful, and salicylates should be employed only to treat erythromelalgia. Oral anticoagulants are not routinely indicated and are difficult to assess owing to the artifactual imbalance between the test tube anticoagulant and plasma that occurs when blood from these patients is assayed for prothrombin or partial thromboplastin activity. Asymptomatic hyperuricemia requires no therapy, but allopurinol should be administered to avoid further elevation of the uric acid when chemotherapy is employed to reduce splenomegaly or leukocytosis-associated pruritus. Generalized pruritus intractable to antihistamines can be a major problem in polycythemia vera, and hydroxyurea, interferon (IFN)-a, and psoralens with ultraviolet light in the A range (PUVA) therapy may have some palliative effects. Asymptomatic thrombocytosis requires no therapy. Symptomatic thrombocytosis or splenomegaly can be treated with hydroxyurea or IFN-a, although each can be associated with significant side effects. Anagrelide, a quinazolin derivative and platelet antiaggregant that also lowers the platelet count, can control thrombocytosis. A reduction in platelet number may be necessary in the treatment of erythromelalgia if salicylates are not effective or if the thrombocytosis is associated with migraine-like symptoms. However, the highest priority for treatment is reduction of the red cell mass to normal. Alkylating agents and ^{32}P are leukemogenic in polycythemia vera, and their use should be avoided. If a cytotoxic agent must be used, hydroxyurea is preferred, but it also may be leukemogenic with chronic use. Chemotherapy should be used for as short a time as possible. In some patients, massive splenomegaly unresponsive to reduction by hydroxyurea or IFN-a therapy and associated with intractable weight loss will require splenectomy. Allogeneic bone marrow transplantation may be effective in young patients.

Patients with polycythemia vera can be expected to live long and useful lives when their red cell mass is effectively managed with phlebotomy. Chemotherapy is never indicated to control the red cell mass unless venous access is impossible.

IDIOPATHIC MYELOFIBROSIS

Idiopathic myelofibrosis (other designations include *agnogenic myeloid metaplasia* or *myelofibrosis with myeloid metaplasia*) is a clonal disorder of a multipotent hematopoietic progenitor cell of unknown etiology characterized by marrow fibrosis, myeloid metaplasia with extramedullary hematopoiesis, and splenomegaly. Idiopathic myelofibrosis is uncommon; in the absence of a specific clonal marker, establishing this

diagnosis is difficult because myelofibrosis and myeloid metaplasia with splenomegaly are also features of both polycythemia vera andCML. Furthermore, myelofibrosis and splenomegaly occur in a variety of benign and malignant disorders (Table 110-3), many of which are amenable to specific therapies not effective in idiopathic myelofibrosis. In contrast to the other chronic myeloproliferative disorders and so-called acute or malignant myelofibrosis, which can occur at any age, idiopathic myelofibrosis primarily afflicts individuals in their sixth decade or later.

ETIOLOGY

The etiology of idiopathic myelofibrosis is unknown. Although nonrandom chromosome abnormalities such as 20q-, 13q-, and trisomy 1q are not uncommon, no specific cytogenetic abnormality has been identified. The degree of myelofibrosis and the extent of extramedullary hematopoiesis are not related. This disorder is associated with overproduction of type III collagen, a finding that has been attributed to platelet-derived growth factor or transforming growth factorb, but no proof has been forthcoming. Importantly, fibroblasts in idiopathic myelofibrosis are not part of the neoplastic clone.

CLINICAL FEATURES

No specific signs or symptoms are associated with idiopathic myelofibrosis. Most patients are asymptomatic at presentation and are usually detected by the discovery of splenic enlargement and/or abnormal blood counts during a routine examination. A blood smear reveals the characteristic features of extramedullary hematopoiesis: teardrop-shaped red cells, nucleated red cells, myelocytes, and promyelocytes; myeloblasts may also be present but have no prognostic significance. Anemia, usually mild initially, is the rule, while the leukocyte and platelet counts are either normal or increased but either can be depressed. Mild hepatomegaly may accompany the splenomegaly, and both the lactate dehydrogenase and serum alkaline phosphatase levels can be elevated. The level of leukocyte alkaline phosphatase can be low, normal, or elevated. Marrow may be unaspirable due to the myelofibrosis, and bone x-rays may reveal osteosclerosis. Exuberant extramedullary hematopoiesis can cause ascites, pulmonary hypertension, intestinal or ureteral obstruction, intracranial hypertension, pericardial tamponade, spinal cord compression, or skin nodules. Splenic enlargement can be sufficiently rapid to cause splenic infarctions with fever and pleuritic chest pain. Hyperuricemia and secondary gout may ensue.

DIAGNOSIS

While the clinical picture described above is characteristic of idiopathic myelofibrosis, all of the clinical features described can be observed in polycythemia vera orCML. Massive splenomegaly commonly masks erythrocytosis in polycythemia vera, and reports of intraabdominal thromboses in idiopathic myelofibrosis likely represent instances of unrecognized polycythemia vera. Furthermore, many other disorders have features that overlap with idiopathic myelofibrosis but respond to distinctly different therapies. Therefore, the diagnosis of idiopathic myelofibrosis is one of exclusion, which requires that the disorders listed inTable 110-3 be ruled out.

The presence of teardrop-shaped red cells, nucleated red cells, myelocytes, and

promyelocytes establishes the presence of extramedullary hematopoiesis; the presence of leukocytosis, thrombocytosis with large and bizarre platelets, as well as circulating myeloblasts suggests the presence of a myeloproliferative disorder as opposed to a secondary form of myelofibrosis (Table 110-3). Marrow is usually not aspirable due to increased marrow reticulin, but marrow biopsy will reveal a hypercellular marrow with trilineage hyperplasia and, in particular, increased megakaryocytes, but there are no characteristic morphologic abnormalities that distinguish idiopathic myelofibrosis from the other chronic myeloproliferative disorders. Splenomegaly due to extramedullary hematopoiesis may be sufficiently massive to cause portal hypertension and variceal formation. In some patients, exuberant extramedullary hematopoiesis can dominate the clinical picture. An intriguing feature of idiopathic myelofibrosis is the occurrence of autoimmune abnormalities such as immune complexes, antinuclear antibodies, rheumatoid factor, or a positive Coombs' test. Whether these represent a host reaction to the disorder or are involved in its pathogenesis is unknown. Cytogenetic analysis of blood or marrow is useful both to exclude CML and for prognostic purposes, because complex karyotype abnormalities portend a poor prognosis in idiopathic myelofibrosis.

COMPLICATIONS

Idiopathic myelofibrosis is a chronic disorder but with a median survival of only 5 years (range 1 to 15 years), a duration much shorter than for polycythemia vera or essential thrombocytosis. The natural history of idiopathic myelofibrosis is one of inexorable marrow failure with transfusion-dependent anemia and increasing organomegaly. Patients are prone to deep-seated tissue infections, particularly of the lungs. As with CML, idiopathic myelofibrosis can evolve from a chronic phase to an accelerated phase with constitutional symptoms and increasing marrow failure. About 10% of patients develop an aggressive form of acute leukemia for which therapy is usually ineffective. Important prognostic factors for disease acceleration include anemia; thrombocytopenia; age; the presence of complex cytogenetic abnormalities; and constitutional symptoms such as unexplained fever, night sweats, or weight loss. Any nonrandom cytogenetic abnormality is associated with a shortened life span, and the presence or development of multiple cytogenetic abnormalities is highly indicative of disease acceleration.

TREATMENT

There is no specific therapy for idiopathic myelofibrosis. Anemia may be exacerbated by deficiency of folic acid or iron, and in rare instances, pyridoxine therapy has been effective. However, anemia is more often due to ineffective erythropoiesis not compensated for by the extramedullary hematopoiesis in the spleen and liver; neither androgens nor erythropoietin has been consistently effective therapy. Erythropoietin may worsen splenomegaly. A red cell splenic sequestration study can establish the presence of hypersplenism, for which splenectomy is indicated. Splenectomy may also be necessary if splenomegaly impairs alimentation and should be performed before cachexia sets in. In this situation, splenectomy should not be avoided because of concern over rebound thrombocytosis, loss of hematopoietic capacity, or compensatory hepatomegaly. However, for unexplained reasons, splenectomy increases the risk of blastic transformation. Allopurinol can control significant hyperuricemia and hydroxyurea has proved useful for controlling organomegaly. The role of interferon-a is undefined,

and its side effects are more pronounced in the older individuals who are affected with this disorder, but reversal of myelofibrosis has been observed. Glucocorticoids are used to control autoimmune complications. Allogeneic bone marrow transplantation should be considered in younger patients.

ESSENTIAL THROMBOCYTOSIS

Essential thrombocytosis (other designations include *essential thrombocythemia*, *idiopathic thrombocytosis*, *primary thrombocytosis*, *hemorrhagic thrombocythemia*) is a clonal disorder of unknown etiology involving a multipotent hematopoietic progenitor cell and is manifested clinically by the overproduction of platelets without a definable cause. Essential thrombocytosis is an uncommon disorder, but its exact frequency is unknown. No clonal marker distinguishes it from the more common nonclonal, reactive forms of thrombocytosis (Table 110-4). Clinical recognition of thrombocytosis is unlikely in the largely asymptomatic persons affected by this disorder. As a consequence, essential thrombocytosis was formerly considered to be a disease of the elderly and to be responsible for significant morbidity due to hemorrhage or thrombosis. However, with the widespread application of platelet counting, it is now clear that essential thrombocytosis can occur at any age in adults and often occurs without symptoms or disturbances of hemostasis. There is an unexplained female predominance, in contrast to the reactive forms of thrombocytosis where no sex bias exists. Because no clonal marker is available for the disorder, clinical criteria have been proposed to distinguish it from the other chronic myeloproliferative disorders, which may also present with thrombocytosis but have distinct prognosis and treatment (Table 110-5). These criteria do not establish clonality; therefore, they are truly useful only in identifying disorders such asCML, polycythemia vera, or myelodysplasia, which can masquerade as essential thrombocytosis, as opposed to establishing the presence of essential thrombocytosis. Furthermore, as with "primary" erythrocytosis, nonclonal, benign forms of thrombocytosis exist (such as hereditary overproduction of thrombopoietin) that are not widely recognized because we currently lack the diagnostic tools to do so.

ETIOLOGY

Megakaryocytopoiesis and platelet production depend upon thrombopoietin and its receptor, Mpl. As in the case of early erythroid and myeloid progenitor cells, early megakaryocytic progenitors require the presence of interleukin (IL) 3 and stem cell factor for optimal proliferation, and their subsequent development is enhanced by IL-6 and -11. However, megakaryocyte maturation and differentiation require thrombopoietin.

Megakaryocytes are unique amongst hematopoietic progenitor cells because they undergo endomitotic as opposed to mitotic reduplication of their genome. In the absence of thrombopoietin, endomitotic megakaryocytic reduplication and, by extension, the cytoplasmic development necessary for platelet production are impaired. Like erythropoietin, thrombopoietin is produced in both the liver and the kidneys, and an inverse correlation between the platelet count and plasma thrombopoietic activity exists. Like erythropoietin, plasma levels of thrombopoietin are controlled in part by the size of its progenitor cell pool. In contrast to erythropoietin, but like its myeloid counterparts granulocyte and granulocyte-macrophage colony stimulating factors, thrombopoietin not only enhances the proliferation of its target cells but also enhances the reactivity of their

end-stage product, the platelet. In addition to its role in thrombopoiesis, thrombopoietin enhances the survival of multipotent hematopoietic stem cells.

The clonality of essential thrombocytosis has been established by the use of the isoenzymes of glucose-6-phosphate dehydrogenase in patients who are hemizygous for this gene, by the use of X-linked DNA polymorphisms, and by the identification of nonrandom, although variable cytogenetic abnormalities. The multipotent hematopoietic progenitor cell involved in this disorder can vary; in some patients lymphocytes contained the same clonal marker as the megakaryocytes, erythrocytes, and myeloid cells, whereas in others the lymphocytes were not involved. Similar observations have been made in polycythemia vera. Furthermore, a number of families have been described in which essential thrombocytosis was inherited, in one instance as an autosomal dominant trait. In one kindred, in addition to essential thrombocytosis, idiopathic myelofibrosis and polycythemia vera were also individually documented.

CLINICAL FEATURES

Clinically, essential thrombocytosis is most often identified incidentally when a platelet count is obtained during the course of a routine evaluation. Occasionally, review of previous platelet counts will reveal that an elevation was present but overlooked. No symptoms or signs are specific for essential thrombocytosis, but patients do have hemorrhagic and thrombotic tendencies expressed as easy bruising for the former or microvascular occlusions for the latter, which may be manifested by erythromelalgia, migraine, or transient ischemic attacks. Physical examination is generally unremarkable except for the presence of mild splenomegaly. Massive splenomegaly is more characteristic of the other myeloproliferative disorders, particularly polycythemia vera or idiopathic myelofibrosis.

Anemia is unusual, but a mild neutrophilic leukocytosis is not. The blood smear, however, is most remarkable for the number of platelets present, some of which may be very large. The leukocyte alkaline phosphatase score is either normal or elevated. The large mass of circulating platelets may prevent the accurate measurement of serum potassium due to the release of platelet potassium upon blood clotting. This hyperkalemia is a laboratory artifact and is not associated with any electrocardiographic abnormalities. Similarly, arterial oxygen measurements can be inaccurate unless the blood is collected on ice. The prothrombin and partial thromboplastin times are normal, while abnormalities of platelet function such as a prolonged bleeding time and impaired platelet aggregation can be present. However, in spite of much study, characteristic platelet function abnormalities associated are not defined, and no platelet function test predicts the presence of clinically significant bleeding or thrombosis.

The elevated platelet count may hinder the collection of a marrow aspirate, but marrow biopsy usually reveals both megakaryocyte hyperplasia and hypertrophy, as well as an overall increase in marrow cellularity. An increase in marrow reticulin may be present, but if extensive, another diagnosis should be considered. The absence of stainable iron demands an explanation, because iron deficiency alone can cause thrombocytosis and absent marrow iron is a feature of polycythemia vera.

While nonrandom cytogenetic abnormalities have been identified in essential

thrombocytosis, no consistently identifiable abnormality is noted, even involving chromosomes 3 and 1 where the genes for thrombopoietin and its receptor Mpl, respectively, are located.

DIAGNOSIS

Thrombocytosis is encountered in a variety of clinical disorders (Table 110-4) in which production of cytokines is increased. Thus, the first obligation when confronted with a high platelet count is to determine if it is a consequence of another disorder. Cytogenetic evaluation is mandatory to determine if the thrombocytosis is due toCML or a myelodysplastic disorder such as the 5q-syndrome. Because the bcr-abl translocation can be present in the absence of the Ph chromosome, polymerase chain reaction analysis for bcr-abl expression should be performed in all patients with thrombocytosis in whom a cytogenetic study is normal. Anemia and ringed sideroblasts are not features of essential thrombocytosis, but they are features of idiopathic refractory sideroblastic anemia, in which thrombocytosis can also occur. The presence of massive splenomegaly should suggest the possibility of another myeloproliferative disorder, and in this setting a red cell mass determination is mandatory because substantial splenomegaly can mask the presence of erythrocytosis. What appears to be essential thrombocytosis can evolve into polycythemia vera, revealing the true nature of the underlying myeloproliferative disorder.

COMPLICATIONS

Perhaps no other condition in clinical medicine has caused otherwise astute physicians to intervene inappropriately more often than thrombocytosis, particularly if the platelet count is greater than 1 ´10^6/uL. It is commonly believed that a high platelet count must cause intravascular stasis and thrombosis; however, no controlled clinical study has ever established either association.

To the contrary, very high platelet counts are associated primarily with hemorrhage, while platelet counts of <1´10^6/uL are more often associated with thrombosis. This is not meant to imply that an elevated platelet count cannot cause symptoms in a patient with essential thrombocytosis, but rather that the focus should be on the patient, not the platelet count. For example, some of the most dramatic neurologic problems in essential thrombocytosis are migraine-related but may respond only to lowering of the platelet count; other symptoms may be a manifestation of erythromelalgia and respond simply to platelet cyclooxygenase inhibitors such as aspirin, without a reduction in platelet number. Still others may represent an interaction between an atherosclerotic vascular system and a high platelet count, and others may have no relationship to the platelet count whatsoever. Progress in distinguishing essential thrombocytosis from polycythemia vera and in defining new causes of hypercoagulability (like factor V Leiden) make the older literature on thrombocytosis less reliable.

TREATMENT

An elevated platelet count in an asymptomatic patient requires no therapy, and before any therapy is initiated in a patient with thrombocytosis, the cause of symptoms must be clearly identified to be a consequence of the elevated platelet count. Plasmapheresis

and cytotoxic therapy have never been proven efficacious and cannot be recommended. Furthermore, patients with essential thrombocytosis treated with ^{32}P, hydroxyurea, or alkylating agents are placed at risk of developed acute leukemia without any proof of benefit from such therapy. If platelet reduction is deemed necessary on the basis of neurologic symptoms refractory to salicylates, IFN-aor anagrelide, a quinazolin derivative, can reduce the platelet count, but neither is uniformly effective nor without significant side effects. Bleeding associated with thrombocytosis usually responds toe-aminocaproic acid, which can be given prophylactically before and after elective surgery. As more clinical experience is acquired, it appears that essential thrombocytosis is more benign than previously thought, and that evolution to acute leukemia is more likely to be a consequence of prior therapy than of the disease itself. In managing patients with thrombocytosis, the physician's first obligation is to do no harm.

(Bibliography omitted in Palm version)

Back to Table of Contents

111. ACUTE AND CHRONIC MYELOID LEUKEMIA - *Meir Wetzler*, *John C. Byrd*, *Clara D. Bloomfield*

The myeloid leukemias are a heterogeneous group of diseases characterized by infiltration of the blood, bone marrow, and other tissues by neoplastic cells of the hematopoietic system. In 2000, the estimated number of new myeloid leukemia cases in the United States was 10,100. These leukemias comprise a spectrum of malignancies that, untreated, range from rapidly fatal to slowly growing. Based on their untreated course, the myeloid leukemias have traditionally been designated *acute* or *chronic*.

ACUTE MYELOID LEUKEMIA

INCIDENCE

The incidence of acute myeloid leukemia (AML) is approximately 2.3 per 100,000 people per year, and the age-adjusted incidence is higher in men than in women (2.9 versus 1.9). AML incidence increases with age; it is 1.3 in individuals younger than 65 years and 12.2 in those older than 65. No significant change in AML incidence has occurred over the past 20 years.

ETIOLOGY

Heredity, radiation, chemical and other occupational exposures, and drugs have been implicated in the development ofAML. No direct evidence suggests a viral etiology.

Heredity Certain syndromes with somatic cell chromosome aneuploidy, e.g., Down (chromosome 21 trisomy), Klinefelter (XXY and variants), and Patau (chromosome 13 trisomy), are associated with an increased incidence ofAML. Inherited diseases with excessive chromatin fragility, e.g., Fanconi anemia, Bloom syndrome, ataxia telangiectasia, and Kostmann syndrome, are also associated with AML.

Radiation Survivors of the atomic bomb explosions in Japan had an increased incidence of myeloid leukemias that peaked 5 to 7 years after exposure. Therapeutic radiation alone seems to add little risk of AML but can increase the risk in people exposed to alkylating agents (see below).

Chemical and Other Exposures Exposure to benzene, which is used as a solvent in the chemical, plastic, rubber, and pharmaceutical industries, is associated with an increased incidence of AML. Smoking and exposure to petroleum products, paint, embalming fluids, ethylene oxide, herbicides, and pesticides, have also been associated with an increased risk of AML.

Drugs Anticancer drugs are the leading cause of treatment-associatedAML. Alkylating agent-associated leukemias occur on average 4 to 6 years after exposure, and affected individuals have aberrations in chromosomes 5 and 7. Topoisomerase II inhibitor-associated leukemias occur 1 to 3 years after exposure, and affected individuals usually have aberrations involving chromosome 11q23. Chloramphenicol, phenylbutazone, and, less commonly, chloroquine and methoxypsoralen can result in bone marrow failure that may evolve into AML.

CLASSIFICATION

The categorization of acute leukemia into biologically distinct groups is based on morphology, cytochemistry, and immunophenotype as well as cytogenetic and molecular techniques.

Morphologic and Cytochemical Classification The diagnosis ofAML is established by the presence of >20% myeloblasts in blood and/or bone marrow. Myeloblasts have nuclear chromatin that is uniformly fine or lacelike in appearance and large nucleoli (two to five per cell). If specific cytoplasmic granules, Auer rods, or the nuclear folding and clefting characteristic of monocytoid cells are not present, the morphologic features observed under light microscopy may not be sufficient to clarify the diagnosis. A positive myeloperoxidase reaction in >3% of the blasts may be the only feature distinguishing AML from acute lymphoblastic leukemia (ALL).

AMLis classified based on morphology and cytochemistry according to the French, American, and British (FAB) schema, which includes eight major subtypes, M0 to M7 (Table 111-1). The World Health Organization classification incorporates molecular (including cytogenetic), morphologic, and clinical features (such as prior hematologic disorder) in defining disease entities (Table 111-1).

Immunophenotypic Classification The phenotype of human myeloid leukemia cells can be studied by multiparameter flow cytometry after the cells are labeled with monoclonal antibodies to cell-surface antigens. For example, M0, which is characterized by immature morphology and no lineage-specific cytochemical reactions, is diagnosed by flow cytometric demonstration of the myeloid-specific antigens cluster designation (CD) 13 or 33. Similarly, M7 can often be diagnosed only by expression of the platelet-specific antigen CD41 or by electron-microscopic demonstration of myeloperoxidase.

Chromosomal Classification Chromosomal analysis of the leukemic cell provides the most important pretreatment prognostic information inAML. Only two cytogenetic abnormalities have been invariably associated with a specificFABgroup: t(15;17)(q22;q12) with M3 and inv(16)(p13q22) with M4Eo. However, many chromosomal abnormalities have been associated primarily with one FAB group, including t(8;21)(q22;q22) with M2, and t(9;11)(p22;q23), and other translocations involving 11q23, with M5. Many of the recurring chromosomal abnormalities in AML have been associated with specific clinical characteristics. More commonly associated with younger age are t(8;21) and t(15;17), and with older age, del(5q) and del(7q). Granulocytic sarcomas are associated with t(8;21); disseminated intravascular coagulation (DIC) with t(15;17); and diabetes insipidus, fever, and infection all with monosomy 7. The reasons for most associations of chromosomal abnormalities with specific clinical features are unknown.

Molecular Classification The many recurring cytogenetic abnormalities led to molecular studies, which have revealed genes that may be involved in leukemogenesis. The 15;17 translocation, characteristic of M3, encodes a chimeric protein, Pml/Rara, which is formed by the fusion of the retinoic acid receptor-a(*RAR*a) gene from

chromosome 17 and the promyelocytic leukemia (*PML*) gene from chromosome 15. The *RAR*a gene encodes a member of the nuclear hormone receptor family of transcription factors. After binding retinoic acid, *RAR*a can promote expression of a variety of genes. The 15;17 translocation juxtaposes *PML* with *RAR*a in a head-to-tail configuration that is under the transcriptional control of *PML*. Three different breakpoints in the *PML* gene lead to various fusion proteins. The Pml-Rara fusion protein tends to suppress gene transcription and blocks differentiation of the cells. Pharmacologic doses of the Rara ligand, all-trans-retinoic acid (tretinoin or ATRA), relieve the block and promote differentiation (see below).

The inv(16), characteristic of M4Eo orAML with abnormal bone marrow eosinophils, and the t(8;21) characteristic of M2 both involve subunits of the transcription factor complex core-binding factor (Cbf), also known as polyomavirus enhancer binding protein 2. This transcription factor contains two subunits, an a subunit, the Am11 protein, and a b subunit, the Pebp2 protein, and is involved in the expression of a number of differentiation-dependent genes in myeloid cells. The inv(16) results in a fusion of the core-binding factorb (*CBFB*) gene on the q arm (encodes Pebp2 protein) and the myosin heavy chain (*MYH11*) gene on the p arm. The 8;21 translocation involves the core-binding factor a(*CBFA*) gene on chromosome 21, called the *AML1* gene, joining the *ETO* gene on chromosome 8. Similar to the t(15;17) gene product, the Aml1/Eto protein acts to block transcription of *CBFA*-*CBFB*-controlled genes.

Most translocations that involve 11q23 rearrange the *MLL* (myeloid-lymphoid or mixed-lineage leukemia) gene. The *MLL* gene has two regions that encompass multiple zinc fingers and has at least two additional potential DNA-binding motifs. Abnormalities in the *MLL* gene are relatively common in patients with AMLwho do not have 11q23 rearrangements cytogenetically.

These molecular aberrations are increasingly being used for diagnosis and detection of residual disease after treatment.

CLINICAL PRESENTATION

Symptoms Patients withAML most often present with nonspecific symptoms that begin gradually or abruptly and are the consequence of anemia, leukocytosis, leukopenia or leukocyte dysfunction, or thrombocytopenia. Nearly half have had symptoms for 3 months or more before the leukemia is diagnosed.

Half mention fatigue as the first symptom, but most complain of fatigue or weakness at the time of diagnosis. Anorexia and weight loss are common. Fever with or without an identifiable infection is the initial symptom in ~10% of patients. Signs of abnormal hemostasis (bleeding, easy bruising) are noted first in 5% of patients. On occasion, bone pain, lymphadenopathy, nonspecific cough, headache, or diaphoresis is the presenting symptom.

Rarely patients may present with symptoms from a mass lesion located in the soft tissues, breast, uterus, ovary, cranial or spinal dura, gastrointestinal tract, lung, mediastinum, prostate, bone, or other organs. The mass lesion represents a tumor of leukemic cells and is called a *granulocytic sarcoma*, or *chloroma*. TypicalAML may

occur simultaneously, later, or not at all in these patients. This rare presentation is more common in patients with 8;21 translocations.

Physical Findings Fever, splenomegaly, hepatomegaly, lymphadenopathy, sternal tenderness, and evidence of infection and hemorrhage are often found at diagnosis. Significant gastrointestinal bleeding, intrapulmonary hemorrhage, or intracranial hemorrhage occur most often in M3AML. Bleeding associated with coagulopathies may also occur in M5 AML and with extreme degrees of leukocytosis or thrombocytopenia in otherFABsubtypes. Retinal hemorrhages are detected in 15% of patients. Infiltration of the gingivae, skin, soft tissues, or the meninges with leukemic blasts at diagnosis is characteristic of the monocytic subtypes (M4 and M5).

Hematologic Findings Anemia is usually present at diagnosis and can be severe. The degree varies considerably irrespective of other hematologic findings, splenomegaly, or the duration of symptoms. The anemia is usually normochromic normocytic. Decreased erythropoiesis often results in a reduced reticulocyte count, and erythrocyte survival is decreased by accelerated destruction. Active blood loss also contributes to the anemia.

The median presenting leukocyte count is about 15,000/uL. Twenty-five to 40% of patients have counts <5000/uL, and 20% have counts>100,000/uL. Fewer than 5% have no detectable leukemic cells in the blood. Poor neutrophil function may be noted functionally by impaired phagocytosis and migration and morphologically by abnormal lobulation and deficient granulation.

Platelet counts<100,000/uL are found at diagnosis in ~75% of patients, and about 25% have counts <25,000/uL. Both morphologic and functional platelet abnormalities can be observed, including large and bizarre shapes with abnormal granulation and inability of platelets to aggregate or adhere normally to one another.

Pretreatment Evaluation Once the diagnosis ofAML is suspected, a rapid evaluation and initiation of appropriate therapy should follow (Table 111-2). In addition to clarifying the subtype of leukemia, initial studies should evaluate the overall functional integrity of the major organ systems, including the cardiovascular, pulmonary, hepatic, and renal systems. Factors that have prognostic significance, either for achieving complete remission (CR) or for predicting the duration of CR, should also be assessed before initiating treatment. Leukemic cells should be obtained from all patients and cryopreserved for future use as new tests become available. All patients should be evaluated for infection.

Most patients are anemic and thrombocytopenic at presentation. Replacement of the appropriate blood components, if necessary, should begin promptly. Because qualitative platelet dysfunction or the presence of an infection may increase the likelihood of bleeding, evidence of hemorrhage justifies the immediate use of platelet transfusion, even if the platelet count is only moderately decreased.

About 50% of patients have a mild to moderate elevation of serum uric acid at presentation. Only 10% have marked elevations, but renal precipitation of uric acid and the nephropathy that may result is a serious but uncommon complication. The initiation of chemotherapy may aggravate hyperuricemia, and patients are usually immediately

started on allopurinol and hydration at diagnosis. Finally, the presence in high concentrations of lysozyme, a marker for monocytic differentiation, may be etiologic in renal tubular dysfunction, which could worsen other renal problems that arise during the initial phases of therapy.

PROGNOSTIC FACTORS

The single most important prognostic factor is attainment ofCR. CR is defined after examination of both blood and bone marrow and should last³4 weeks. The blood neutrophil count must be³1500/uL and the platelet count³100,000/uL. Hemoglobin concentration or hematocrit are not considered in determining CR. Circulating blasts should be absent. While rare blasts may be detected in the blood during marrow regeneration, they should disappear on successive studies. Bone marrow cellularity should be >20% with trilineage maturation. The bone marrow should contain <5% blasts, and Auer rods should be absent. Extramedullary leukemia should not be present. For patients in CR, reverse transcriptase polymerase chain reaction (RT-PCR) to detectAML-associated molecular abnormalities, and fluorescence in situ hybridization (FISH) to detect AML-associated cytogenetic aberrations are currently used to detect residual disease. Such detection of minimal residual disease may become a reliable discriminator between patients in CR who do or do not require additional and/or alternative therapies.

Many factors influence the likelihood of enteringCR, the length of CR, and the curability ofAML. Prognostic factors are influenced by the treatment used. Age at diagnosis remains among the most important pretreatment risk factors, with>60 years being associated with a poorer prognosis primarily because of its influence on the patient's ability to survive induction therapy and thus achieve CR. Chronic and intercurrent diseases impair tolerance to rigorous therapy; acute medical problems at diagnosis reduce the likelihood of survival. Performance status, independent of age, also influences ability to survive induction therapy and thus respond to treatment. Age may also influence outcome because AML in older patients differs biologically. The leukemic cells in elderly patients more commonly express CD34 and the mdr1 efflux pump that conveys resistance to natural product-derived agents such as the anthracyclines (see below). With each successive decade of age, a greater proportion of patients have more resistant disease.

Chromosome findings at diagnosis are an independent prognostic factor. Patients with t(8;21), inv(16), or t(15;17) have extremely good prognoses, while those with no cytogenetic abnormality have a moderately favorable outcome when treated with high-dose cytarabine. Patients with del(5q), -7, and abnormalities involving 12p have a very poor prognosis. Patients with certain abnormalities, such as inv(3), rarely achieveCR with standard induction chemotherapy.

A prolonged symptomatic interval with cytopenias preceding diagnosis or a history of an antecedent hematologic disorder are other pretreatment clinical features that are associated with a lowerCR rate and shorter survival time. The CR rate is lower in patients who have had anemia, leukopenia, and/or thrombocytopenia for >1 month before the diagnosis of AML when compared to those without such a history. Responsiveness to chemotherapy declines as the duration of the antecedent disorder(s)

increases. SecondaryAMLdeveloping after treatment with cytotoxic agents and/or irradiation for other malignancies is extremely difficult to treat successfully.

A high presenting leukocyte count is an independent prognostic factor; duration ofCR is inversely related to the presenting leukocyte count or absolute circulating myeloblast count. Among patients with hyperleukocytosis (>100,000/uL), early central nervous system bleeding and pulmonary leukostasis and late relapse contribute to poor outcome.

TheFABclassification diagnosis has been found to be an independent prognostic factor in some series. Other characteristics of the leukemic cell have been reported to have prognostic significance, including Auer rods, ultrastructural features, in vitro and in vivo growth characteristics and chemotherapeutic sensitivity, and immunophenotype. Expression of the *MDR1* gene adversely influences outcome. This gene encodes a protein that actively pumps out a variety of lipophilic compounds (e.g., anthracyclines) from the cell.

In addition to pretreatment variables, several treatment factors have been reported to correlate with prognosis inAML, in particular withCRduration. One is the rapidity with which the blast cells disappear from the blood after the institution of therapy. In addition, patients who achieve CR after one induction cycle have longer CR than those requiring multiple cycles.

TREATMENT

Treatment of the newly diagnosed patient withAML is usually divided into two phases, induction and postremission management (Fig. 111-1). The initial goal is to quickly induceCR. Once CR is obtained, further therapy must be used to prolong survival.

Induction Chemotherapy The most commonly usedCRinduction regimens (for patients with allFABsubtypes except M3) consist of combination chemotherapy with cytarabine (cytosine arabinoside) and an anthracycline. Cytarabine is a cell cycle S-phase-specific antimetabolite that becomes phosphorylated to an active triphosphate form that interferes with DNA synthesis. Anthracyclines are DNA intercalaters. Their primary mode of action is thought to be inhibition of topoisomerase II, leading to DNA breaks. Cytarabine is usually administered as a continuous intravenous infusion at 100 to 200 mg/m^{2}per day for 7 days. Anthracycline therapy generally consists of daunorubicin, 45 mg/m^{2}intravenously on days 1, 2, and 3 (*the 7 and 3 regimen*). Treatment with idarubicin at 12 or 13 mg/m^{2}per day for 3 days in conjunction with cytarabine by 7-day continuous infusion is at least as effective and may be superior to daunorubicin in younger patients. The addition of etoposide or other agents does not increase the CR rate but may improve the CR duration.

After induction chemotherapy, the bone marrow is examined to determine if the leukemia has been eliminated. If >5% blasts exist with^{3}20% cellularity, the patient has traditionally been retreated with cytarabine and an anthracycline in doses similar to those given initially, but for 5 and 2 days, respectively. Our recommendation, however, is to consider changing therapy in this setting. Patients who fail to attainCR after two induction courses should immediately proceed to an allogeneic stem cell transplant

(SCT) if an appropriate donor exists.

With the 7 and 3 cytarabine/daunorubicin regimen outlined above, 65 to 75% of adults with de novoAMLachieveCR. Two-thirds achieve CR after a single course of therapy, and one-third require two courses. About 50% of patients who do not achieve CR have a drug-resistant leukemia, and 50% do not achieve CR because of fatal complications of bone marrow aplasia or impaired recovery of normal stem cells.

High-dose cytarabine-based regimens have very highCR rates after a single cycle of therapy. When given in high doses, more cytarabine may enter the cells, saturate the cytarabine-inactivating enzymes, and increase the intracellular levels of 1-b-D-arabinofuranylcytosine-triphosphate, the active metabolite incorporated into DNA. Thus, higher doses of cytarabine may increase the inhibition of DNA synthesis and thereby overcome resistance to standard-dose cytarabine. In two randomized studies, one by the Southwest Oncology Group (SWOG) and one by the Australian Leukemia Study Group (ALSG), high-dose cytarabine with an anthracycline produced CR rates similar to those achieved with standard 7 and 3 regimens. However, the ALSG demonstrated that the CR duration was much longer after high-dose cytarabine than after standard-dose cytarabine.

The hematologic toxicity of high-dose cytarabine-based induction regimens has typically been greater than that associated with 7 and 3 regimens. Toxicity with high-dose cytarabine includes myelosuppression, pulmonary toxicity, and significant and occasionally irreversible cerebellar toxicity. All patients treated with high-dose cytarabine must be closely monitored for cerebellar toxicity. Full cerebellar testing should be performed before each dose, and further high-dose cytarabine should be withheld if evidence of cerebellar toxicity develops.

Supportive Care Measures geared to supporting patients through several weeks of granulocytopenia and thrombocytopenia are critical to the success ofAMLtherapy. Patients with AML should be treated in centers expert in providing supportive measures for their management.

Recombinant hematopoietic growth factors have been incorporated into clinical trials inAML. These trials have been designed to lower the infection rate after chemotherapy or to sensitize (prime) the leukemic blasts to chemotherapy, or both. Both granulocyte colony stimulating factor (G-CSF) and granulocyte-macrophage colony stimulating factor (GM-CSF) have reduced the median time to neutrophil recovery by an average of 5 to 7 days. This accelerated rate of neutrophil recovery, however, has not always translated into significant reductions in infection rates. In most randomized studies, both G-CSF and GM-CSF have failed to improve the CR rate, disease-free survival, or overall survival. Although receptors for both G-CSF and GM-CSF are present on AML blasts, therapeutic efficacy is neither enhanced nor inhibited by these agents. The use of growth factors as supportive care for AML patients is controversial. We favor their use in elderly patients, those receiving intensive regimens, patients with uncontrolled infections, or those participating in clinical trials.

Multilumen right atrial catheters should be inserted through a subcutaneous tunnel as soon as patients with newly diagnosedAML have been stabilized. They should be used

thereafter for administration of intravenous medications and transfusions, as well as for blood drawing. The separation between the vascular access site and the exit site and the presence of a Dacron cuff in the subcutaneous channel reduce the risk of infection. With meticulous attention to sterile technique in catheter placement and maintenance, catheters may often be left in place for months.

Adequate and prompt blood bank support is critical to therapy ofAML. Platelet transfusions should be given as needed to maintain a platelet count>10,000 to 20,000/uL. We believe that the platelet count should be kept at higher levels in febrile patients and during episodes of active bleeding orDIC. Patients with poor posttransfusion platelet count increments may benefit from administration of platelets from human leukocyte antigen (HLA)-matched donors. Red blood cell transfusions should be administered to keep the hemoglobin level>80 g/L (8 g/dL) in the absence of active bleeding or DIC. Blood products leukodepleted by filtration should be used to avert or delay alloimmunization as well as febrile reactions. Blood products should also be irradiated to prevent graft-versus-host disease (GVHD). Cytomegalovirus (CMV)-negative blood products should be used for CMV-seronegative patients who are potential candidates for allogeneicSCT. Leukodepleted products are also effective for these patients if CMV-negative products are not available.

Infectious complications remain the major cause of morbidity and death during induction and postremission chemotherapy forAML. Prophylactic administration of antibiotics in the absence of fever is controversial. Oral nystatin or clotrimazole are recommended to prevent localized candidiasis. For patients who are herpes simplex virus antibody titer-positive, acyclovir prophylaxis is effective in preventing reactivation of latent oral herpes infections.

Fever develops in most patients withAML, but infections are documented in only half of febrile patients. Early initiation of empiric broad-spectrum antibacterial and antifungal antibiotics has significantly reduced the number of patients dying of infectious complications (Chap. 85). An antibiotic regimen adequate to treat gram-negative and gram-positive organisms should be instituted at the onset of fever in a granulocytopenic patient after clinical evaluation, including a detailed physical examination with inspection of the indwelling catheter exit site and a perirectal examination, as well as procurement of cultures and radiographs aimed at documenting the source of fever. Specific antibiotic regimens should be based on antibiotic sensitivity data obtained from the institution at which the patient is being treated. Acceptable regimens include imipenem-cilastin, an antipseudomonal semisynthetic penicillin (e.g., piperacillin) combined with an aminoglycoside, a third-generation cephalosporin with antipseudomonal activity (i.e., ceftazidine or cefapime) or doubleb-lactam combinations (ceftazidine and piperacillin). Empiric vancomycin is not given initially in the absence of suspected gram-positive infection or mucositis. Aminoglycosides should be avoided if possible in patients with renal insufficiency. For patients with known immediate-type hypersensitivity reactions to penicillin, aztreonam may be substituted for b-lactams. Aztreonam should be combined with an aminoglycoside or a quinolone antibiotic rather than used alone. Empiric vancomycin should be initiated in neutropenic patients who remain febrile for 3 days, and amphotericin B is added at 7 days if fever persists. Liposomal amphotericin is at least equivalent to regular amphotericin for empiric antifungal treatment and has less renal toxicity. Antibacterial and antifungal antibiotics

should be continued until patients are no longer neutropenic, regardless of whether a specific source has been found for the fever.

Treatment of Promyelocytic Leukemia ATRA is an oral drug that induces the differentiation of leukemic cells bearing the t(15;17); it is not effective in other forms ofAML. Acute promyelocytic leukemia is responsive to cytarabine and daunarubicin, but about 10% of patients treated with these drugs die fromDICinduced by the release of granule components by dying tumor cells. ATRA does not produce DIC but produces another complication called the retinoic acid syndrome. Occurring within the first 3 weeks of treatment, it is characterized by fever, dyspnea, chest pain, pulmonary infiltrates, pleural and pericardial effusions, and hypoxia. The syndrome is related to the adhesion of differentiated neoplastic cells in the pulmonary vasculature. Glucocorticoids, chemotherapy, and/or supportive measures can be effective. About 10% of patients die from this syndrome.

ATRA (45 mg/$m_2$per day orally until remission is documented) plus concurrent chemotherapy (7 and 3) appears to be the safest and most effective treatment for acute promyelocytic leukemia. Unlike patients with other types ofAML, patients with this subtype may benefit from maintenance therapy with either ATRA or chemotherapy. The optimal regimen is being sought in clinical studies.

Arsenic trioxide produces meaningful responses in patients refractory to ATRA.

The detection of minimal residual disease byRT-PCRamplification of the t(15;17) chimeric gene product appears to predict relapse. Disappearance of the signal is associated with long-term disease-free survival; its persistence predicts relapse. With increases in the sensitivity of the assay, some patients with persistent abnormal gene product have been found who do not suffer a relapse. It is not known whether there is a critical threshold level of transcripts that predicts for leukemia relapse.

Postremission Therapy Induction of a durable firstCR is critical to long-term disease-free survival inAML. Without further therapy virtually all patients experience relapse. Once relapse has occurred, AML is generally curable only bySCT.

Postremission therapy is designed to eradicate any residual leukemic cells. Therefore, it should prevent relapse and prolong survival. Approaches to postremission therapy in AMLinclude intensive chemotherapy and allogeneic or autologousSCT. In patients older than 65 years, such therapy is of uncertain benefit. High-dose cytarabine is more effective than standard-dose cytarabine. The Cancer and Leukemia Group B, for example, compared the duration ofCR in patients randomly assigned postremission to four cycles of high (3 g/$m_2$every 12 h on days 1, 3, and 5), intermediate (400 mg/$m_2$for 5 days by continuous infusion), or standard (100 mg/$m_2$per day for 5 days by continuous infusion) doses of cytarabine. A dose-response effect for cytarabine in patients with AML who were age 60 years or younger was demonstrated. High-dose cytarabine significantly prolonged CR and increased the fraction cured in patients with favorable [t(8;21) and inv(16)] and normal cytogenetics, but it had no significant effect on patients with other abnormal karyotypes.

Allogeneic and autologousSCT in first CRhas been studied extensively in younger

patients with no major organ dysfunction. Allogeneic SCT is used in patients <55 years with an HLA-compatible donor. Relapse with this therapy occurs in only a small fraction of patients, but toxicity is relatively high from treatment; complications include veno-occlusive disease,GVHD, and infections. Autologous transplantation can be used in young and older patients and uses the same type of high-dose therapy. Patients subsequently receive their own stem cells collected while in remission. The toxicity is lower with autologous SCT (5% mortality rate), but the relapse rate is higher than with allogeneic SCT. The increased relapse rate is due to the absence of the graft-vs-leukemia effect seen with allogeneic SCT and possible contamination of the autologous stem cells with tumor cells. Purging the autologous stem cells does not lower the relapse rate with autologous SCT.

Randomized trials comparing intensive therapy and autologous and allogeneicSCT have shown improved duration of remission with allogeneic SCT compared to autologous SCT or chemotherapy alone. However, overall survival is generally not different; the improved disease control with allogeneic SCT is erased by the increase in fatal toxicity. Prognostic factors may help select patients in firstCR for whom transplant is most effective.

Our approach includes strong consideration for allogeneicSCT in first CRfor patients with high risk karyotypes. Patients with normal karyotypes who have other poor risk factors (antecedent hematologic disorder, failure to attain remission with a single induction course, hyperleukocytosis,*MLL*gene abnormalities) are also potential candidates. If a suitable HLA donor does not exist, autologous SCT or novel therapeutic approaches are considered. Patients with t(8;21) and inv(16) are treated with repetitive doses of high-dose cytarabine, which offers a high frequency of cure without the morbidity of transplant.

Relapse Once relapse occurs after the standard induction and postremission chemotherapy approach described above and outlined in Fig. 111-1, patients are rarely cured with further standard-dose chemotherapy. Patients eligible for allogeneicSCTshould receive transplants expeditiously at the first sign of relapse. Long-term disease-free survival is approximately the same (30 to 50%) with allogeneic SCT in first relapse or in second remission. Autologous SCT rescues about 20% of relapsed patients with AML who have chemosensitive disease. The most important factors predicting response at relapse are the length of the previousCR, whether initial CR was achieved with one or two courses of chemotherapy, and the type of postremission therapy. Because of the poor outcome of patients in early (<12 months) first relapse, it is justified (for patients without HLA-compatible donors) to explore innovative approaches, such as new drugs or immunotherapies (Table 111-3). Patients with longer (>12 months) first CR generally relapse with drug-sensitive disease and may achieve a second remission with the original induction regimen. However, cure for these patients is uncommon, and treatment with novel approaches should be considered if SCT is not possible. It is not yet clear whether careful monitoring for residual disease byRT-PCR,FISH and quantitative PCR identifies patients destined to relapse who are more readily cured by salvage therapy given before overt clinical relapse.

CHRONIC MYELOID LEUKEMIA

INCIDENCE

The incidence of chronic myeloid leukemia (CML) is 1.3 per 100,000 people per year, and the age-adjusted incidence is higher in men than in women (1.7 versus 1.0). CML incidence decreased slightly between 1973 and 1991 (1.5 versus 1.3). The incidence of CML increases slowly with age until the middle forties, when it starts to rise rapidly.

DEFINITION

The diagnosis ofCML is established by identifying a clonal expansion of a hematopoietic stem cell possessing a reciprocal translocation between chromosomes 9 and 22. This translocation results in the head-to-tail fusion of the breakpoint cluster region (*BCR*) gene on chromosome 22q11 with the *ABL* (named after the abelson murine leukemia virus) gene located on chromosome 9q34. Untreated, the disease is characterized by the inevitable transition from a chronic phase to an accelerated phase and on to blast crisis.

ETIOLOGY

No clear correlation with exposure to cytotoxic drugs, such as alkylating agents, has been found, and there is no direct evidence of a viral etiology. Cigarette smoking has been shown to accelerate the progression to blast crisis and therefore has an adverse effect on survival inCML. The effect of radiation was demonstrated in the study of the atomic bomb survivors, where it has been estimated that the development of a CML cell mass of 10,000/uL takes 6.3 years. No increase in CML incidence was found in the survivors of the Chernobyl accident, suggesting that only large doses of radiation can induce CML.

PATHOPHYSIOLOGY

The product of the fusion gene resulting from the t(9;22) plays a central role in the development ofCML. This chimeric gene is transcribed into a hybrid*BCR*/*ABL* mRNA in which exon 1 of ABL is replaced by variable numbers of 5¢ BCR exons. Bcr/Abl fusion proteins, p210$_{BCR\text{-}ABL}$, are produced that contain NH_2-terminal domains of Bcr and the COOH-terminal domains of Abl. Bcr/Abl fusion proteins can transform hematopoietic progenitor cells in vitro. Furthermore, reconstituting lethally irradiated mice with bone marrow cells infected with retrovirus carrying the gene encoding the p210$_{BCR\text{-}ABL}$leads to the development of a myeloproliferative syndrome resembling CML in 50% of the mice. Specific antisense oligomers to the *BCR/ABL* junctions inhibit the growth of t(9;22)-positive leukemic cells without affecting normal colony formation.

The mechanism(s) by which p210$_{BCR\text{-}ABL}$promotes the transition from the benign state to the fully malignant one is still unclear. Messenger RNA for *BCR/ABL* can occasionally be detected in normal individuals. However, attachment of the*BCR*sequences to *ABL* results in three critical functional changes: (1) the Abl protein becomes constitutively active as a tyrosine kinase enzyme, (2) the DNA protein-binding activity of Abl is attenuated, and (3) the binding of Abl to cytoskeletal actin microfilaments is enhanced.

Disease Progression The events associated with transition to the acute phase are

poorly understood. Chromosomal instability of the malignant clone, resulting, for example, in the acquisition of an additional t(9;22), trisomy 8, or 17p- (p53 loss), is a fundamental characteristic ofCML. Acquisition of these additional genetic and/or molecular abnormalities is critical to the phenotypic transformation. The site of the breakpoint within the*BCR*gene may predict the time to development of blast crisis, but this claim has been refuted by others. Heterogeneous structural alterations of the p53 gene, as well as structural alterations and lack of protein production of the retinoblastoma gene, have been associated with disease progression in a subset of patients. Rare patients show alterations in *RAS*. Sporadic reports also document the presence of an altered *MYC* (named after the myelocytomatosis virus) gene or the appearance of p190*BCR-ABL*, the protein commonly found in adultALL and occasionally inAML, during the clinical evolution of small numbers of patients with CML. Progressive de novo DNA methylation at the *BCR/ABL* locus has also been shown to herald blastic transformation. Finally, interleukin (IL)-1b may be involved in the progression of CML to the blastic phase. Multiple pathways to disease transformation exist, but the exact timing and relevance of each of these remains unclear.

CLINICAL PRESENTATION

Symptoms The clinical onset of the chronic phase is generally insidious. Accordingly, some patients are diagnosed while still asymptomatic, during health screening tests; other patients present with fatigue, malaise, and weight loss or have symptoms resulting from splenic enlargement, such as early satiety and left upper quadrant pain or mass. Less common are features related to granulocyte or platelet dysfunction, such as infections, thrombosis, or bleeding. Occasionally, patients present with leukostatic manifestations due to severe leukocytosis or thrombosis such as vasoocclusive disease, cerebrovascular accidents, myocardial infarction, venous thrombosis, priapism, visual disturbances, and pulmonary insufficiency.

Progression ofCML is associated with worsening symptoms. Unexplained fever, significant weight loss, increasing dose requirement of the drugs controlling the disease, bone and joint pain, bleeding, thrombosis, and infections suggest transformation into accelerated or blastic phases. Fewer than 10 to 15% of newly diagnosed patients present with accelerated disease or with de novo blastic phase CML.

Physical Findings In most patients the abnormal finding on physical examination at diagnosis is minimal to moderate splenomegaly; mild hepatomegaly is found occasionally. Persistent splenomegaly despite continued therapy is a sign of disease acceleration. Lymphadenopathy and extramedullary myeloid tumors (granulocytic sarcomas) are unusual except late in the course of the disease; when they are present, the prognosis is poor.

Hematologic Findings Elevated white blood cell counts, with various degrees of immaturity of the granulocytic series, are present at diagnosis. Usually <5% circulating blasts and <10% blasts and promyelocytes are noted. Cycling of the counts may be observed in patients followed without treatment. Platelet counts are almost always elevated at diagnosis, and a mild degree of normochromic normocytic anemia is present. Leukocyte alkaline phosphatase is characteristically low inCMLcells. Serum levels of vitamin B_{12} and vitamin B_{12}-binding proteins are generally elevated. Phagocytic

functions are usually normal at diagnosis and remain normal during the chronic phase. Histamine production secondary to basophilia is increased in later stages, causing pruritus, diarrhea, and flushing.

At diagnosis, bone marrow cellularity, primarily of the myeloid and megakaryocytic lineages, with a greatly altered myeloid to erythroid ratio, is increased in almost all patients withCML. The marrow blast percentage is generally normal or slightly elevated. Marrow or blood basophilia, eosinophilia, and monocytosis may be present. While collagen fibrosis in the marrow is unusual at presentation, significant degrees of reticulin stain-measured fibrosis are noted in about half of the patients.

Disease acceleration is defined by the development of increasing degrees of anemia unaccounted for by bleeding or chemotherapy, cytogenetic clonal evolution, or blood or marrow blasts between 10 and 20%, blood or marrow basophils³20%, or platelet count<100,000/uL. *Blast crisis* is defined as acute leukemia, with blood or marrow blasts ³20%. Hyposegmented neutrophils may appear (Pelger-Huet anomaly). Blast cells can be classified as myeloid, lymphoid, erythroid, or undifferentiated, based on morphologic, cytochemical, and immunologic features. About half the cases are myeloid, one-third lymphoid, 10% erythroid, and the rest are undifferentiated.

Chromosomal Findings The cytogenetic hallmark ofCML, found in 90 to 95% of patients, is the t(9;22)(q34;q11). Originally, this was recognized by the presence of a shortened chromosome 22 (22q-), designated as the *Philadelphia chromosome*, that arises from the reciprocal 9;22 translocation. Some patients may have complex translocations (designated as *variant translocations*) involving three, four, or five chromosomes (usually including chromosomes 9 and 22). However, the molecular consequences of these changes appear similar to those resulting from the typical t(9;22).

PROGNOSTIC FACTORS

The clinical outcome of patients with CMLis variable. Death is expected in 10% of patients within 2 years and in about 20% yearly thereafter. The median survival time is ~4 years. Therefore, several prognostic models that identify different risk groups in CML have been developed. The most commonly used staging systems have been derived from multivariate analyses of prognostic factors. The Sokal index identified percentage of circulating blasts, spleen size, platelet count, cytogenetic clonal evolution, and age as the most important prognostic indicators. Two models, that of Tura and the combined model of Kantarjian, divide patients according to the number of negative prognostic factors. Age³60 years, spleen³10 cm below the costal margin, blasts³3% in blood or³5% in marrow, basophils³7% in blood or³3% in marrow, platelets³700,000/uL, or any of the characteristics of accelerated disease are associated with a very poor short-term prognosis and a threefold higher hazard rate, or risk of death per unit of time, in the first year. A prognostic scoring system to estimate the survival of CML patients treated with interferon (IFN) a has been developed.

TREATMENT

The goal of therapy inCML is to achieve prolonged, durable, nonneoplastic, nonclonal

hematopoiesis, which entails the eradication of any residual cells containing theBCR/ABL transcript. Hence the goal is complete molecular remission and cure (Table 111-4). A proposed treatment plan for the newly diagnosed patient with CML is presented inFig. 111-2.

Allogeneic SCT AllogeneicSCT is the only curative therapy forCML and, when feasible, is the treatment of choice. However, it is complicated by a high early mortality rate owing to the transplant procedure. When the outcome of all patients undergoing allogeneic SCT reported to the International Bone Marrow Transplant Registry was compared with the outcome of all patients treated with hydroxyurea or interferon (IFN) a by the German CML Study Group, the survival for the former group was statistically better, but only starting 5 years after transplant. When only low-risk patients (by Sokal's criteria) were evaluated, the benefit in survival for allogeneic SCT was seen after 6 years. Outcome of SCT depends on multiple factors including: (1) the patient (i.e., age and phase of disease); (2) the type of donor [i.e., syngeneic (monozygotic twins) or HLA-compatible allogeneic, related or unrelated]; (3) the preparative regimen; (4)GVHD; and (5) posttransplantation treatment.

The Patient As experience has been gained and safety and efficacy have been established, it has become clear that patients should be younger than 65 years and have a healthy and histocompatible donor. Furthermore, survival afterSCT in the accelerated and blastic phases of the disease is significantly diminished and is associated with a very high rate of relapse. The Seattle data demonstrate that SCT early in the chronic phase (1 to 2 years from diagnosis) is superior to later SCT. While overall survival, disease-free survival, and relapse rates are not influenced by priorIFN-atreatment, incidence and severity of acute and chronicGVHDcorrelate with prior IFN-a treatment in the unrelated donor and possibly also in the related donor setting. Therefore, because early SCT is more effective than late SCT, the decision to perform allogeneic SCT should probably be made within a year of diagnosis when IFN-a is the initial therapy; a 3-month hiatus is recommended between discontinuing IFN-a and initiating SCT.

The Donor Transplantation from a family donor, who is either fully matched or mismatched at only one HLA locus, should be considered standard therapy for any patient withCML who is a candidate for an HLA-related sibling transplant. SyngeneicSCT in patients with chronic phase CML has been reported from the Seattle group to result in 7-year disease-free survival in 55%, with a 30% relapse rate. With HLA-identical sibling SCT in the chronic phase, many groups have reported 5-year disease-free survival in 40 to 70% of patients, with a 25% relapse rate. SCT from an HLA-matched unrelated donor has been reported by the Seattle group to result in a 74% probability of surviving 3 years for patients transplanted in chronic phase less than 1 year from diagnosis and younger than 50 years. A 2-year disease-free (based on hematologic analyses) survival of 45%(±21%) for patients receiving SCT from unrelated individuals, matched or mismatched at only one locus, transplanted less than 1 year from diagnosis was reported by the National Marrow Donor Program. Patients age 40 to 50 years fared poorly (15 of 55 survived). The probability of 2-year disease-free (based on cytogenetic analyses) survival in a report on unrelated transplants from the Medical College of Wisconsin, using a standardized conditioning regimen and T cell depletion, was 52% for patients transplanted in chronic phase (regardless of the time from

diagnosis). Patients receiving transplants from unrelated individuals have higher rates of graft failure and acute and chronicGVHD and prolonged convalescence after treatment, compared to those who receive allogeneic transplants from related individuals. Peripheral blood is now being studied as a source of hematopoietic progenitor cells; it may offer rapid engraftment and less risk for the donor. Umbilical-cord blood may permit mismatched SCT with notably less GVHD; graft-versus-leukemia (GVL) effects do not appear to be impaired. A problem with cord blood is obtaining an appropriate number of progenitor cells to reconstitute hematopoiesis in an adult.

Preparative Regimens These regimens have been studied by several groups. A randomized study by the Seattle group compared cyclophosphamide and total-body irradiation with busulphan and cyclophosphamide. They found no significant differences in the 3-year probabilities of survival, relapse, event-free survival, speed of engraftment, or incidence of venoocclusive disease of the liver. Significantly more patients in the total-body irradiation arm experienced major elevations of creatinine, acuteGVHD, longer periods of fever, positive blood cultures, hospital admissions, and longer inpatient hospital stays. However, increased chronic GVHD, obstructive bronchiolitis and alopecia were noted with busulphan. Intravenous busulphan may permit better control of serum levels. Minitransplants in which the preparative regimen is aimed at eliminating host lymphocytes rather than bone marrow are being tested. Reduced toxicity with preserved antitumor efficacy is the goal.

Development and type of GVHD Development of grade I GVHD, as compared to no GVHD, decreases the risk of relapse. A lower relapse rate was observed also in patients with grade II GVHD but was accompanied by a substantially higher transplant-related mortality rate. The decreased relapse rate may be caused by aGVLeffect. Depletion of T lymphocytes from donor marrow can prevent GVHD but results in an increased risk of relapse, which exceeds the relapse rate after syngeneicSCT. Thus, T lymphocytes from the donor marrow mediate a significant antileukemic, or GVL, effect, and even syngeneic marrow may exhibit limited GVL activity inCML.

Posttransplantation Treatment Further support for the existence of an immunologically mediatedGVLeffect comes from the observation that donor leukocyte infusions (without prior conditioning orGVHDprophylaxis) can induce hematologic and cytogenetic remissions in patients withCML who have relapsed after allogeneicSCT.

The activity ofIFN-ain patients with early chronic-phaseCML was the basis for the use of IFN-a afterSCT, either to induce cytogenetic remissions in relapsed patients or to prevent relapse after SCT for high-risk patients. The main concern about IFN-a use after allogeneic SCT has been the development or worsening ofGVHD, because IFN-a acts as an immunomodulator. However, in published reports encompassing 52 allogeneic recipients who were free of GVHD and were either at high risk for relapse or had already relapsed, only 6 subsequently developed GVHD after IFN-a therapy was initiated. IFN-a has also been combined with mononuclear cells obtained from donor blood to induce cytogenetic remissions in relapsed patients. Cytogenetic remissions have been achieved, but the exact role of IFN-a as opposed to the mononuclear cells is unclear.IL-2, with or without IFN-a, is also being evaluated for its ability to restore complete cytogenetic remission in patients who suffer relapses after SCT.

IFN-ahas been used afterSCT to prevent relapse in patients with advanced disease at time of transplant (patients at high risk for relapse). CytogeneticCR has been maintained for as long as 2 years posttransplantation in small numbers of patients with blast crisis or second chronic phase. Similarly, IL-2 (2.5´10^6 to 6´10^6units/m^2per day) has been given to patients after T cell-depleted allogeneic SCT in an effort to induceGVLwithoutGVHD, and thus prevent relapse. Compared with historical control subjects, patients treated with IL-2 have a lower risk of disease relapse. A randomized trial is warranted.

Interferons When allogeneicSCT is not feasible,IFN-atherapy is the treatment of choice. The interferons are a complex group of naturally occurring proteins produced by eukaryotic cells in response to viruses, antigens, and mitogens. Three distinct groups of IFN species have been identified: IFN-a, -b, and -g. Although various interferons have become available for clinical investigation, most data have been generated with IFN-apreparations.

Interferons have potent, pleiotropic biologic effects, spanning a spectrum of antiviral, microbicidal, immunomodulatory, and antiproliferative properties. While interferons downregulate the expression of several oncogenes and cytokines, they also upregulate the expression ofIFNregulatory factor-1 (a transcriptional activator with antioncogenic activity), adhesion molecules, and the histocompatibility genes. Interferons also inhibit angiogenesis and induce a cellular immune response. However, their mode(s) of action are still unknown.

In seven randomized studies comparingIFN-aand chemotherapy, both modalities have been found to be effective in achieving hematologic remissions. However, patients treated with IFN-a survived longer than patients treated with hydroxyurea or busulphan. The 5-year survival rate was 51% with IFN-a and 42% with chemotherapy.

Patients develop both acute and chronic side effects fromIFN-atherapy. Acute side effects (flulike symptoms) appear early in the course of the treatment. Most flulike symptoms respond to acetaminophen, and tachyphylaxis develops within 1 to 2 weeks. Chronic reactions, such as fatigue and lethargy, depression, weight loss, myalgias, and arthralgias, occur in about half of the patients and may require dose reduction. Patients also report cough, postnasal drip, and dryness of the skin. Infrequently, immune-mediated thrombocytopenia and anemia develop. In addition, long-term therapy has been associated with late autoimmune side effects, such as hypothyroidism and occasionally generalized autoimmune phenomena.

The most important persistent side effects in patients withCML who are treated withIFN-aare neurologic. All patients treated with IFN-a are subject to some neurologic toxicity, the most common symptom being lethargy. Up to 20% of patients have neurologic side effects that are associated with compromised quality of life and reduced ability to carry out their regular activity, such as full-time work. In addition, at the required doses, impotence in men is not infrequent.

Hematologic remissions are generally achieved within 1 to 2 months of startingIFN-a. However, some patients have a cyclic response pattern with progressively lower peak

and nadir counts over a period of months. The increase in counts during the cycling that occurs in the first few months of therapy should not be confused with resistance. Cytogenetic responses generally start at 3 to 12 months, and complete cytogenetic responses may require 6 months to 4 years of therapy. However, most complete cytogenetic responses are achieved within 12 to 18 months; and in single-agent, single-arm studies they have been identified in up to 26% of patients.

The combination ofIFN-awith cytarabine has produced better results than those with IFN-a alone; cytogenetic responses occurred earlier, but the influence on survival is not yet known.

Chemotherapy Initial management of patients with chemotherapy is currently reserved for rapid lowering of white blood cell counts, reduction of symptoms, and reversal of symptomatic splenomegaly. Hydroxyurea, a ribonucleotide reductase inhibitor, induces rapid disease control. The initial dose is 1 to 4 g/d, and the dose should be reduced by half with each 50% reduction of the leukocyte count. Unfortunately, cytogenetic remissions with hydroxyurea are uncommon. Busulphan, an alkylating agent that acts on early progenitor cells, has a more prolonged effect. However, we do not recommend its use because of its serious side effects, which include unexpected, and occasionally fatal, myelosuppression in 5 to 10% of patients; pulmonary, endocardial, and marrow fibrosis; and an Addison-like wasting syndrome.

Homoharringtonine (HHT) is a plant alkaloid derived from a tree, *Cephalotaxus fortuneii* sp. *harringtonii*. HHT blocks peptide bond formation after binding of the aminoacyl-transfer RNA to the ribosome. In patients whose disease progressed during treatment withIFN-aor who were in later chronic phase (>1 year from diagnosis), HHT induced 72% complete hematologic responses and 22% complete or major cytogenetic responses. The use of HHT before IFN-a in early chronic phase resulted in a 92% complete hematologic response rate and a 27% major cytogenetic response rate. Toxicity is mainly related to myelosuppression.

Intensive combination chemotherapy has also been used in chronic phaseCML, with 30 to 50% of patients achieving complete cytogenetic responses. However these cytogenetic remissions have been short lived. Consequently, intensive combination chemotherapy regimens are being used today only to mobilize normal progenitors in the blood in order to collect circulating stem cells for autologous transplantation.

Autologous SCT Autologous SCT could potentially cureCML if a means to select the residual normal progenitors, which coexist with their malignant counterparts, could be developed. As a source of autologous hematopoietic stem cells for transplantation, blood offers certain advantages over marrow (e.g., faster engraftment and no general anesthesia). Normal hematopoietic stem cells appear with increased frequency in the blood of patients with CML during the recovery phase after chemotherapy andG-CSF.

A retrospective analysis of >200 autologousSCTperformed forCML at eight centers worldwide suggests that autologous SCT prolongs survival in chronic- or accelerated-phase patients when compared with conventional therapy. At transplant 93 patients were in chronic phase, 25 were in accelerated phase, and 114 were in blast crisis or second chronic phase. Patients received autologous bone marrow and/or blood

hematopoietic stem cells. In 42 cases the hematopoietic progenitors were subjected to ex vivo manipulation by long-term bone marrow culture, by incubation with recombinantIFN-g, or by chemotherapy. In 49 cases the hematopoietic progenitors were harvested during the recovery phase after various chemotherapy regimens. After autologous SCT, 29 of 93 (31%) patients in first chronic phase achieved complete cytogenetic remissions. The median duration of cytogenetic remission was 14 months, with a range of 2 to 68 months. Approaches to treat minimal residual disease after autologous transplantation, such as immune modulation, are currently being investigated.

Leukapheresis and Splenectomy Intensive leukapheresis may control the blood counts in chronic phaseCML; however, it is expensive and cumbersome. It is useful in emergencies where leukostasis-related complications such as pulmonary failure or cerebrovascular accidents are likely. It may also have a role in the treatment of pregnant women in whom it is important to avoid potentially teratogenic drugs.

Splenectomy was used inCML in the past because of the suggestion that evolution to the acute phase might occur in the spleen. However, this does not appear to be the case, and splenectomy is now reserved for symptomatic relief of painful splenomegaly unresponsive to chemotherapy or for significant anemia or thrombocytopenia associated with hypersplenism. Splenic radiation is used rarely to reduce the size of the spleen.

Minimal Residual Disease The correlation between residual cells with the t(9;22) and disease recurrence is not completely understood. In initial studies withRT-PCRused to predict disease recurrence afterIFN-atherapy, residual disease was found in all samples tested from patients with complete cytogenetic remissions. Later studies demonstrated the elimination of theBCR/ABL mRNA transcript after more prolonged IFN-a treatment in some cases. It is now possible to quantitate transcripts, and longer follow-up may indicate whether quantitation of the BCR/ABL transcript is useful for predicting cytogenetic and clinical relapse.

After allogeneicSCT,RT-PCRanalysis may be positive for residual disease during the first 6 months in patients who subsequently achieve a long-lasting remission. However, late persistence of RT-PCR positivity appears to indicate a reduced probability of cure. RT-PCR positivity at any single time point is not predictive of imminent relapse. After allogeneic SCT, patients are often divided according to RT-PCR results into one of three groups: (1) persistently positive, (2) intermittently negative, and (3) persistently negative. These three groups have low, intermediate, and high probability of maintaining remission and disease free-survival, respectively. Although these data suggest that patients who are persistently RT-PCR positive more than 6 months after allogeneic SCT need additional therapeutic interventions, this conclusion has not been rigorously established. The studies have used an assortment of techniques for measuring minimal residual disease, the level of sensitivity has been variable, and the follow-up durations of patients are short. Real-time RT-PCR may provide a more sensitive tool to predict relapse inCML and in other cancers. In patients who do not have any evidence forGVHD and are intermittently RT-PCR negative,GVL may be induced by alloreactive donor cells (without the side effects of GVHD) to suppress the proliferation of the leukemic cells.

Future Directions The synthetic inhibitor of theBCR/ABL kinase, STI571, induces selective inhibition in the growth of t(9;22)-bearing tumor cells in vitro and some responses in patients. Inhibition of *RAS* with a farnesyl transferase inhibitor that blocks its insertion into the membrane may have antitumor activity inCML on the basis of early clinical trials. Preclinical efforts to use BCR/ABL peptides as a tumor vaccine appear promising. The use of BCR/ABL antisense oligonucleotides to purge residual leukemic cells from autologous hematopoietic progenitors before reinfusion, as well as approaches to induceGVL in the setting of minimal residual disease without inducingGVHD, are underway.

Treatment of Blast Crisis The treatment for all forms of blast crisis is generally ineffective. Treatment is tailored to the phenotype of the blast cell. Myeloid crises and erythroid crisis are treated as forAML, but remissions occur in only a minority of cases and are generally short lived. Patients may present without having had a chronic phase. AML with a t(9;22) is probably blast crisis ofCML and carries a poor prognosis.

Lymphoid blast crisis is treated likeALL(Chap. 112) with vincristine (1.4 mg/m^2weekly) plus prednisone (60 mg/m^2orally qd) induction therapy with or without an anthracycline. About one-third of patients will reenter chronic phase after 2 to 3 weeks of treatment, but the remissions last only a median of ~4 months. EvenSCT is minimally effective during blast crises. Novel treatment approaches are needed.

(Bibliography omitted in Palm version)

Back to Table of Contents

112. MALIGNANCIES OF LYMPHOID CELLS - *James O. Armitage*, *Dan L. Longo*

Malignancies of lymphoid cells range from the most indolent to the most aggressive human malignancies. These cancers arise from cells of the immune system at different stages of differentiation, resulting in a wide range of morphologic, immunologic, and clinical findings. Advances in our understanding of the normal immune system have allowed a better understanding of these sometimes confusing disorders.

Some malignancies of lymphoid cells almost always present as leukemia (i.e., primary involvement of bone marrow and blood), while others almost always present as lymphomas (i.e., solid tumors of the immune system). However, other malignancies of lymphoid cells can present as either leukemia or lymphoma. In addition, the clinical pattern can change over the course of the illness. This change is more often seen in a patient who seems to have a lymphoma and then develops the manifestations of leukemia over the course of the illness.

BIOLOGY OF LYMPHOID MALIGNANCIES: CONCEPTS OF THE WHO CLASSIFICATION OF LYMPHOID MALIGNANCIES

The classification of lymphoid malignancy evolved steadily throughout the twentieth century. The distinction between leukemia and lymphoma was made early, and separate classification systems were developed for each. Leukemias were first divided into acute and chronic subtypes based on average survival. Chronic leukemias were easily subdivided into those of lymphoid or myeloid origin based on morphologic characteristics. However, in recent years, a spectrum of diseases that were formerly all called chronic lymphoid leukemia has become apparent (Table 112-1). The acute leukemias were usually malignancies of blast cells with few identifying characteristics. When cytochemical stains became available, it was possible to divide these objectively into myeloid malignancies and acute leukemias of lymphoid cells. Acute leukemias of lymphoid cells have been subdivided based on morphologic characteristics by the French-American-British (FAB) group (Table 112-2). Using this system, lymphoid malignancies of small uniform blasts (e.g., typical childhood acute lymphoblastic leukemia) were called L1, lymphoid malignancies with larger and more variable size cells were called L2, and lymphoid malignancies of uniform cells with basophilic and sometimes vacuolated cytoplasm were called L3 (e.g., typical Burkitt's lymphoma cells). Acute leukemias of lymphoid cells have also been subdivided based on immunologic (i.e., T vs. B) and cytogenetic abnormalities (Table 112-2). Major cytogenetic subgroups include the t(9;22) (e.g., Philadelphia chromosome-positive acute lymphoblastic leukemia) and the t(8;14) found in the L3 or Burkitt's leukemia.

Non-Hodgkin's lymphomas were separated from Hodgkin's disease by recognition of the Sternberg-Reed cells early in the twentieth century. The first systematic classification for non-Hodgkin's lymphomas was proposed by Gall and Mallory in the first half of the twentieth century and divided non-Hodgkin's lymphomas into giant follicular lymphoma, lymphosarcoma, and reticulum cell sarcoma. Unfortunately, this fairly simple system proved to be imprecise in its definitions and only marginally clinically useful. In the 1950s, Henry Rappaport and colleagues recognized the importance of growth pattern in subdividing non-Hodgkin's lymphomas and used pattern in addition to cell size and shape as the basis for a new classification that proved more clinically relevant. In the

1970s, it was recognized that non-Hodgkin's lymphomas were all tumors of lymphocytes and were derived from either T or B cells. This led to immunologically based classifications of lymphomas such as the Lukes-Collins classification in the United States and the Kiel classification proposed by Lennert and associates in Europe. In an attempt to unify terminology and improve the effectiveness of communication between pathologists and clinicians, the Working Formulation was proposed in 1982. Over the next two decades the Kiel classification dominated clinical practice in Europe, whereas the Working Formulation became the main classification system used in North America.

In the past two decades, increased understanding of the immune system and the genetic abnormalities associated with non-Hodgkin's lymphoma have led to the identification of several previously unrecognized types of lymphoma. The recognition of these new and clinically relevant lymphomas led to proposals for changing existing classifications. A new proposal that is part of the basis of the new World Health Organization classification of lymphoid malignancies takes into account morphologic, clinical, immunologic, and genetic information and attempts to divide non-Hodgkin's lymphomas and other lymphoid malignancies into clinical/pathological entities that have clinical and therapeutic relevance. This system is presented in Table 112-3. Clinical studies have shown that this new system is clinically relevant and has a higher degree of diagnostic accuracy than those used previously. The possibilities for subdividing lymphoid malignancies are extensive. However,Table 112-3 presents in bold those malignancies that occur in at least 1% of patients. Specific lymphoma subtypes will be dealt with in more detail below.

GENERAL ASPECTS OF LYMPHOID MALIGNANCIES

ETIOLOGY AND EPIDEMIOLOGY

Chronic lymphoid leukemia (CLL) is the most prevalent form of leukemia in western countries. It occurs most frequently in older adults and is exceedingly rare in children. Approximately 13,000 new cases are diagnosed in the United States each year, but because of the prolonged survival associated with this disorder, the total prevalence is many times higher. CLL is more common in men than in women and more common in whites than in blacks. This is an uncommon malignancy in Asia. The etiologic factors for typical CLL are unknown.

In contrast toCLL, acute lymphoid leukemias (ALLs) are predominantly cancers of children and young adults. The L3 or Burkitt's leukemia occurring in children in developing countries seems to be associated with infection by the Epstein-Barr virus (EBV) in infancy. However, the explanation for the etiology of more common subtypes of ALL is much less certain. Childhood ALL occurs more often in higher socioeconomic subgroups. Children with trisomy 21 (Down's syndrome) have an increased risk for childhood acute lymphoblastic leukemia as well as acute myeloid leukemia. Exposure to high-energy radiation in early childhood increases the risk of developing T cell acute lymphoblastic leukemia.

The etiology ofALL in adults is also uncertain. ALL is unusual in middle-aged adults but increases in incidence in the elderly. However, acute myeloid leukemia is still much more common in these patients. Environmental exposures including certain industrial

exposures, exposure to agricultural chemicals, and smoking might increase the risk of developing ALL as an adult.

The cell of origin of Hodgkin's disease has not been determined definitively, but molecular evidence suggests that most are of B cell origin. The incidence of Hodgkin's disease appears fairly stable, with approximately 8000 new cases diagnosed each year in the United States. Hodgkin's disease is more common in whites than in blacks and more common in males than in females. A bimodal distribution of age at diagnosis has been observed, with one peak incidence occurring in patients in their 20s and the other in those in their 80s. Patients in the younger age groups diagnosed in the United States largely have the nodular sclerosing subtype of Hodgkin's disease. Elderly patients, patients infected with HIV, and patients in third world countries more commonly have mixed-cellularity Hodgkin's disease or lymphocyte-depleted Hodgkin's disease. Infection by HIV is a risk factor for developing Hodgkin's disease. In addition, an association between infection byEBV and Hodgkin's disease has been demonstrated. A monoclonal or oligoclonal proliferation of EBV-infected cells in 20 to 40% of the patients with Hodgkin's disease has led to proposals for this virus having an etiologic role in Hodgkin's disease. However, the matter is not settled definitively.

For unknown reasons, non-Hodgkin's lymphomas have increased in frequency in the United States at the rate of 4% per year since 1950. Almost 60,000 new cases of non-Hodgkin's lymphoma were diagnosed in the United States in the year 2000. Non-Hodgkin's lymphomas are more frequent in the elderly and more frequent in men. Patients with both primary and secondary immunodeficiency states are predisposed to developing non-Hodgkin's lymphomas. These include patients with HIV infection; patients who have undergone organ transplantation; and patients with inherited immune deficiencies, the sicca syndrome, and rheumatoid arthritis.

The incidence of non-Hodgkin's lymphomas and the patterns of expression of the various subtypes differ geographically. T cell lymphomas are more common in Asia than in western countries, while certain subtypes of B cell lymphomas such as follicular lymphoma are more common in western countries. A specific subtype of non-Hodgkin's lymphoma known as the angiocentric nasal T/natural killer (NK) cell lymphoma has a striking geographic occurrence, being most frequent in Southern Asia and parts of Latin America. Another subtype of non-Hodgkin's lymphoma associated with infection by human T cell lymphotropic virus (HTLV) I is seen particularly in southern Japan and the Caribbean.

A number of environmental factors have been implicated in the occurrence of non-Hodgkin's lymphoma, including infectious agents, chemical exposures, and medical treatments. Several studies have demonstrated an association between exposure to agricultural chemicals and an increased incidence in non-Hodgkin's lymphoma. Patients treated for Hodgkin's disease can develop non-Hodgkin's lymphoma; it is unclear whether this is a consequence of the Hodgkin's disease or its treatment. However, the infectious etiology of non-Hodgkin's lymphoma is the area where evidence has been expanding most rapidly in recent years. Table 112-4 illustrates those infectious agents associated with the development of non-Hodgkin's lymphoma.HTLV-I infects T cells and leads directly to the development of adult T cell lymphoma (ATL) in a small percentage of infected patients. The cumulative lifetime risk of developing lymphoma in an infected

patient is 2.5%. The virus is transmitted by infected lymphocytes ingested by nursing babies of infected mothers, blood-borne transmission, or sexually. The median age of patients with ATL is about 56 years, emphasizing the long latency.

EBVis associated with the development of Burkitt's lymphoma in Central Africa and the occurrence of aggressive non-Hodgkin's lymphomas in immunosuppressed patients in western countries. EBV infection is strongly associated with the occurrence of extranodal nasal T/NKcell lymphomas in Asia and South America. Infection with HIV predisposes to the development of aggressive, B cell non-Hodgkin's lymphoma. This may be through overexpression of interleukin 6 by infected macrophages. Infection of the stomach by the bacterium *Helicobacter pylori* induces the development of gastric MALT (mucosa-associated lymphoid tissue) lymphomas. This association is supported by evidence that patients treated with antibiotics to eradicate *H. pylori* have regression of their MALT lymphoma. The bacterium does not transform lymphocytes to produce the lymphoma; instead, a vigorous immune response is made to the bacterium and the chronic antigenic stimulation leads to the neoplasia.

Chronic hepatitis C virus infection has been associated with the development of lymphoplasmacytic lymphoma. Human herpesvirus 8 is associated with primary effusion lymphoma in HIV-infected persons and multicentric Castleman's disease, a diffuse lymphadenopathy associated with systemic symptoms of fever, malaise, and weight loss.

In addition to infectious agents, a number of other diseases or exposures may predispose to developing lymphoma (Table 112-5).

IMMUNOLOGY

All lymphoid cells are derived from a common hematopoietic progenitor that gives rise to lymphoid, myeloid, erythroid, monocyte, and megakaryocyte lineages. Through the ordered and sequential activation of a series of transcription factors, the cell first becomes committed to the lymphoid lineage and then gives rise to B and T cells. About 75% of all lymphoid leukemias and 90% of all lymphomas are of B cell origin. A cell becomes committed to B cell development when it begins to rearrange its immunoglobulin genes. The sequence of cellular changes, including changes in cell-surface phenotype, that characterize normal B cell development are shown in Fig. 112-1. A cell becomes committed to T cell differentiation upon migration to the thymus and rearrangement of T cell antigen receptor genes. The sequence of the events that characterize T cell development are depicted in Fig. 112-2.

Although lymphoid malignancies often retain the cell-surface phenotype of lymphoid cells at particular stages of differentiation, this information is of little consequence. The so-called stage of differentiation of a malignant lymphoma does not predict its natural history. For example, the clinically most aggressive lymphoid leukemia is Burkitt's leukemia, which has the phenotype of a mature follicle center IgM-bearing B cell. Leukemias bearing the immunologic cell-surface phenotype of more primitive cells (e.g., pre-BALL, CD10+) are less aggressive and more amenable to curative therapy than the "more mature" appearing Burkitt's leukemia cells. Furthermore, the apparent stage of differentiation of the malignant cell does not reflect the stage at which the genetic

lesions that gave rise to the malignancy developed. For example, follicular lymphoma has the cell-surface phenotype of a follicle center cell, but its characteristic chromosomal translocation, the t(14;18), which involves juxtaposition of the antiapoptotic *bcl-2* gene next to the immunoglobulin heavy chain gene (see below), had to develop early in ontogeny as an error in the process of immunoglobulin gene rearrangement. Why the subsequent steps that led to transformation became manifest in a cell of follicle center differentiation is not clear.

The major value of cell-surface phenotyping is to aid in the differential diagnosis of lymphoid tumors that appear similar by light microscopy. For example, benign follicular hyperplasia may resemble follicular lymphoma; however, the demonstration that all the cells bear the same immunoglobulin light chain isotype strongly suggests the mass is a clonal proliferation rather than a polyclonal response to an exogenous stimulus.

GENETIC CONSIDERATIONS

Malignancies of lymphoid cells are associated with recurring genetic abnormalities. While specific genetic abnormalities have not been identified for all subtypes of lymphoid malignancies, it is presumed that they exist. Genetic abnormalities can be identified at a variety of levels including gross chromosomal changes (i.e., translocations, additions, or deletions); rearrangement of specific genes that may or may not be apparent from cytogenetic studies; and overexpression, underexpression, or mutation of specific oncogenes. Altered expression or mutation of specific proteins is particularly important. Many lymphomas contain balanced chromosomal translocations involving the antigen receptor genes; immunoglobulin genes on chromosomes 2, 14, and 22 in B cells; and T cell antigen receptor genes on chromosomes 7 and 14 in T cells. The rearrangement of chromosome segments to generate mature antigen receptors must create a site of vulnerability to aberrant recombination. B cells are even more susceptible to acquiring mutations during their maturation in germinal centers; the generation of antibody of higher affinity requires the introduction of mutations into the variable region genes in the germinal centers. Other nonimmunoglobulin genes, for example *bcl-6*, may acquire mutations as well.

In the case of diffuse large B cell lymphoma, the translocation t(14;18) occurs in approximately 30% of patients and leads to overexpression of the *bcl-2* gene found on chromosome 18. Some other patients without the translocation also overexpress the BCL-2 protein. This protein is involved in suppressing apoptosis -- i.e., the mechanism of cell death most often induced by cytotoxic chemotherapeutic agents. A higher relapse rate has been observed in patients whose tumors overexpress the BCL-2 protein, but not in those patients whose lymphoma cells show only the translocation. Thus, particular genetic mechanisms have clinical ramifications.

Table 112-6 presents the best documented translocations and associated oncogenes for various subtypes of lymphoid malignancies. In some cases, such as the association of the t(14;18) in follicular lymphoma, the t(2;5) in anaplastic large T/null-cell lymphoma, the t(8;14) in Burkitt's lymphoma, and the t(11;14) in mantle cell lymphoma, the great majority of tumors in patients with these diagnoses display these abnormalities. In other types of lymphoma where a minority of the patients have tumors expressing specific genetic abnormalities, the defects may have prognostic significance. No specific genetic

abnormalities have been identified in Hodgkin's disease.

In typical B cellCLL, trisomy 12 conveys a poorer prognosis. InALL in both adults and children, genetic abnormalities have important prognostic significance. Patients whose tumor cells display the t(9;22) have a much poorer outlook than patients who do not have this translocation. Other genetic abnormalities that occur frequently in adults with ALL include the t(4;11) and the t(8;14). The t(4;11) is associated with younger age, female predominance, high white cell counts, and L1 morphology. The t(8;14) is associated with older age, male predominance, frequent central nervous system (CNS) involvement, and L3 morphology. Both are associated with a poor prognosis. In childhood ALL, hyperdiploidy has been shown to have a favorable prognosis.

Approach to the Patient

Regardless of the type of lymphoid malignancy, the initial evaluation of the patient should include performance of a careful history and physical examination. These will help confirm the diagnosis, identify those manifestations of the disease that might require prompt attention, and aid in the selection of further studies to optimally characterize the patient's status to allow the best choice of therapy. It is difficult to overemphasize the importance of a carefully done history and physical examination. They might provide observations that lead to reconsidering the diagnosis, provide hints at etiology, clarify the stage, and allow the physician to establish rapport with the patient that will make it possible to develop and carry out a therapeutic plan.

For patients withALL, evaluation is usually completed after a complete blood count, chemistry studies reflecting major organ function, a bone marrow biopsy with genetic and immunologic studies, and a lumbar puncture. The latter is necessary to rule out occultCNSinvolvement. At this point, most patients would be ready to begin therapy. In ALL, prognosis is dependent upon the genetic characteristics of the tumor, the patient's age, the white cell count, and the patient's overall clinical status and major organ function.

InCLL, the patient evaluation should include a complete blood count, chemistry tests to measure major organ function, serum protein electrophoresis, and a bone marrow biopsy. However, some physicians believe that the diagnosis would not always require a bone marrow biopsy. Patients often have imaging studies of the chest and abdomen looking for pathologic lymphadenopathy. Patients with typical B cell CLL can be subdivided into three major prognostic groups. Those patients with only blood and bone marrow involvement by leukemia but no lymphadenopathy, organomegaly, or signs of bone marrow failure have the best prognosis. Those with lymphadenopathy and organomegaly have an intermediate prognosis, and patients with bone marrow failure, defined as hemoglobin<100 g/L (10 g/dL) or platelet count<100,000/uL, have the worst prognosis. The pathogenesis of the anemia or thrombocytopenia is important to discern. The prognosis is adversely affected when either or both of these abnormalities are due to progressive marrow infiltration and loss of productive marrow. However, either or both may be due to autoimmune phenomena or to hypersplenism that can develop during the course of the disease. These destructive mechanisms are usually completely reversible (glucocorticoids for autoimmune disease; splenectomy for hypersplenism) and do not influence disease prognosis.

Two popular staging systems have been developed to reflect these prognostic groupings (Table 112-7). Patients with typical B cellCLL can have their course complicated by immunologic abnormalities including autoimmune hemolytic anemia, autoimmune thrombocytopenia, and hypogammaglobulinemia. Patients with hypogammaglobulinemia benefit from regular (monthly)g globulin administration. Because of expense,g globulin is often withheld until the patient experiences a significant infection. These abnormalities do not have a clear prognostic significance and should not be used to assign a higher stage.

The initial evaluation of a patient with Hodgkin's disease or non-Hodgkin's lymphoma is similar. In both situations, the determination of an accurate anatomic stage is an important part of the evaluation. The staging system that is utilized is the Ann Arbor staging system originally developed for Hodgkin's disease (Table 112-8).

Evaluation of patients with Hodgkin's disease will typically include a complete blood count; erythrocyte sedimentation rate; chemistry studies reflecting major organ function; computed tomography (CT) scans of the chest, abdomen, and pelvis; and a bone marrow biopsy. A gallium scan is not necessary for primary staging, but when it is performed at the completion of therapy it allows evaluation of persistent radiographic abnormalities, particularly the mediastinum. In most cases, these studies will allow assignment of anatomic stage and the development of a therapeutic plan.

In patients with non-Hodgkin's lymphoma, the same evaluation described for patients with Hodgkin's disease is usually carried out. In addition, serum levels of lactate dehydrogenase (LDH) andb_2-microglobulin and serum protein electrophoresis are often included in the evaluation. Anatomic stage is assigned in the same manner as used for Hodgkin's disease. However, the prognosis of patients with non-Hodgkin's lymphoma is best assigned using the International Prognostic Index (IPI) (Table 112-9). This is a powerful predictor of outcome in all subtypes of non-Hodgkin's lymphoma. Patients are assigned an IPI score based on the presence or absence of five adverse prognostic factors and may have none or all five of these adverse prognostic factors.Figure 112-3 shows the prognostic significance of this score in 1300 patients with all types of non-Hodgkin's lymphoma.

CLINICAL FEATURES, TREATMENT, AND PROGNOSIS OF SPECIFIC LYMPHOID MALIGNANCIES

PRECURSOR CELL B CELL NEOPLASMS

Precursor B Cell Lymphoblastic Leukemia/Lymphoma The most common cancer in childhood is B cell acute lymphoblastic leukemia (ALL). Although this disorder can also present as a lymphoma in either adults or children, presentation as lymphoma is quite rare.

The malignant cells in patients with precursor B cell lymphoblastic leukemia are most commonly of pre-B cell origin. Patients typically present with signs of bone marrow failure such as pallor, fatigue, bleeding, fever, and infection related to peripheral blood cytopenias. Peripheral blood counts regularly show anemia and thrombocytopenia but

might show leukopenia, a normal leukocyte count, or leukocytosis based largely on the number of circulating malignant cells (seePlate V-24). Extranodal sites of disease are frequently involved in patients who present with leukemia, which might be manifested by lymphadenopathy, hepato- or splenomegaly,CNSdisease, testicular enlargement, and/or cutaneous infiltration.

The diagnosis is usually made by bone marrow biopsy, which shows infiltration by malignant lymphoblasts. Demonstration of a pre-B cell immunophenotype (Fig. 112-1) and, often, characteristic cytogenetic abnormalities (Table 112-6) confirm the diagnosis. An adverse prognosis in patients with precursor B cell ALL is predicted by a very high white cell count, the presence of symptomaticCNSdisease, and unfavorable cytogenetic abnormalities. For example, t(9;22) is frequently found in adults with B cell lymphoblastic leukemia and is associated with a very poor outlook.

TREATMENT

The treatment of patients with precursor B cell lymphoblastic leukemia involves remission induction with combination chemotherapy, a consolidation phase that includes administration of high-dose systemic therapy and treatment to eliminate disease in theCNS, and a period of continuing therapy to prevent relapse and effect cure. The overall cure rate in children is 85%, while about 50% of adults are long-term disease-free survivors. This reflects the high proportion of adverse cytogenetic abnormalities seen in adults with precursor B cell lymphoblastic leukemia.

Precursor B cell lymphoblastic lymphoma is a rare presentation of precursor B cell lymphoblastic malignancy. These patients often have a rapid transformation to leukemia, and similar treatment approaches as are used in patients presenting with leukemia are appropriate. In the few patients who present with the disease confined to lymph nodes, a high cure rate has been reported.

MATURE (PERIPHERAL) B CELL NEOPLASMS

B Cell Chronic Lymphoid Leukemia/Small Lymphocytic Lymphoma B cellCLL/small lymphocytic lymphoma represents by far the most common lymphoid leukemia, and when presenting as a lymphoma, it accounts for ~7% of non-Hodgkin's lymphomas. As the name implies, presentation can be as either leukemia or lymphoma. The major clinical characteristics of B cell CLL/small lymphocytic lymphoma are presented in Table 112-10.

The diagnosis of typical B cellCLL is made when an increased number of circulating lymphocytes (i.e.,>4 ´10^9/L and usually >10 ´10^9/L) is found (see Plate V-17) that are monoclonal B cells and display the CD5 antigen. Confirmation of bone marrow infiltration by the same cells confirms the diagnosis. The peripheral blood smear in such patients typically shows many "smudge" or "basket" cells, nuclear remnants of cells damaged by the physical shear stress of making the blood smear. If cytogenetic studies are performed, trisomy 12 is found in ~25 to 30% of patients. Abnormalities in chromosome 13 are also seen.

If the primary presentation is lymphadenopathy and a lymph node biopsy is performed,

pathologists usually have little difficulty in making the diagnosis of small lymphocytic lymphoma based on morphologic findings and immunophenotype. However, even in these patients ~70 to 75% will be found to have bone marrow involvement and the search for circulating monoclonal B lymphocytes is often positive.

The differential diagnosis of typical B cellCLL is extensive and presented in Table 112-1. Immunophenotyping will eliminate the T cell disorders and can often help sort out other B cell malignancies. For example, only mantle cell lymphoma and typical B cell CLL are usually CD5 positive. Typical B cell small lymphocytic lymphoma can be confused with other B cell disorders including lymphoplasmacytic lymphoma (i.e., the tissue manifestation of Waldenstrom's macroglobulinemia), nodal marginal zone B cell lymphoma, and mantle cell lymphoma. In addition, some small lymphocytic lymphomas have areas of large cells that can lead to confusion with diffuse large B cell lymphoma. An expert hematopathologist is vital for making this distinction.

Typical B cellCLL is often found incidentally when a complete blood count is done for another reason. However, complaints that might lead to the diagnosis include fatigue, frequent infections, and new lymphadenopathy. The diagnosis of typical B cell CLL should be considered in a patient presenting with an autoimmune hemolytic anemia or autoimmune thrombocytopenia. B cell CLL has also been associated with red cell aplasia. When this disorder presents as lymphoma, the most common abnormality is asymptomatic new lymphadenopathy, with or without splenomegaly. The staging systems used to predict prognosis in patients with typical B cell CLL are presented in Table 112-7. TheIPI for non-Hodgkin's lymphomas, which also predicts prognosis in these patients, is presented in Table 112-9. The evaluation of a new patient with typical B cell CLL/small lymphocytic lymphoma will include many of the studies included in Table 112-11, which describes the initial evaluation of a new patient with non-Hodgkin's lymphoma. In addition, particular attention needs to be given to detecting immune abnormalities such as autoimmune hemolytic anemia, autoimmune thrombocytopenia, hypogammaglobulinemia, and red cell aplasia.

TREATMENT

Patients whose presentation is typical B cellCLL with no manifestations of the disease other than bone marrow involvement and lymphocytosis (i.e., Rai stage O and Binet stage A;Table 112-7) can be followed without specific therapy for their malignancy. These patients have a median survival>10 years, and some will never require therapy for this disorder. If the patient has an adequate number of circulating normal blood cells and is asymptomatic, many physicians would not initiate therapy for patients in the intermediate stage of the disease manifested by lymphadenopathy and/or hepatosplenomegaly. However, the median survival for these patients is ~7 years, and most will require treatment in the first few years of follow-up. Patients who present with bone marrow failure (i.e., Rai stage III or IV or Binet stage C) will require initial therapy in almost all cases. These patients have a serious disorder with a median survival of only 1.5 years. It must be remembered that immune manifestations of typical B cell CLL should be managed independently of specific antileukemia therapy. For example, glucocorticoid therapy for autoimmune cytopenias and g globulin replacement for patients with hypogammaglobulinanemia should be used whether or not antileukemia therapy is given.

Patients who present primarily with lymphoma and have a lowIPIscore have a 5-year survival of ~75%, but those with a high IPI score have a 5-year survival of <40% and are more likely to require early therapy.

The most common treatments for patients with typical B cellCLL/small lymphocytic lymphoma have been the use of single-agent chlorambucil or single-agent fludarabine. Chlorambucil can be administered orally with few immediate side effects, while fludarabine is administered intravenously and is associated with significant immune suppression. However, fludarabine is by far the more active agent and is the only drug associated with a significant incidence of complete remission. For young patients presenting with leukemia requiring therapy, fludarabine is today the treatment of choice. Because fludarabine is an effective second-line agent in patients with tumors unresponsive to chlorambucil, the latter agent is often chosen in elderly patients who require therapy. Many patients who present with lymphoma will receive a combination chemotherapy regimen used in other lymphomas such as CVP (cyclophosphamide, vincristine, and prednisone), or CHOP (cyclophosphamide, doxorubicin, vincristine, and prednisone). Young patients with this disease can be candidates for bone marrow transplantation. Allogeneic bone marrow transplantation can be curative but is associated with a significant treatment-related mortality. The place of autologous transplantation in patients with this disorder remains uncertain.

Molecular analysis of immunoglobulin gene sequences inCLL has demonstrated that about half the patients have tumors expressing mutated immunoglobulin genes and half have tumors expressing unmutated or germ-line immunoglobulin sequences. Patients with unmutated immunoglobulins tend to have a more aggressive clinical course and are less responsive to therapy. Unfortunately, immunoglobulin gene sequencing is not routinely available. CD38 expression is said to be low in the better-prognosis patients expressing mutated immunoglobulin and high in poorer-prognosis patients expressing unmutated immunoglobulin, but this test has not been confirmed as a reliable means of distinguishing the two groups.

Extranodal Marginal Zone B Cell Lymphoma of MALT Type Extranodal marginal zone B cell lymphoma ofMALTtype makes up approximately 8% of non-Hodgkin's lymphomas. This small-cell lymphoma presents in extranodal sites. It was previously considered a small lymphocytic lymphoma or sometimes a pseudolymphoma. The recognition that the gastric presentation of this lymphoma was associated with *H. pylori* infection was an important step in recognizing it as a separate entity. The clinical characteristics of extranodal marginal zone B cell lymphoma of MALT type are presented in Table 112-10.

The diagnosis of extranodal marginal zone B cell lymphoma ofMALTtype can be made accurately by an expert hematopathologist based on a characteristic pattern of infiltration of small lymphocytes that are monoclonal B cells and CD5 negative. In some cases, transformation to diffuse large B cell lymphoma occurs, and both diagnoses may be made in the same biopsy. The differential diagnosis includes benign lymphocytic infiltration of extranodal organs and other small-cell B cell lymphomas.

Extranodal marginal zone B cell lymphoma ofMALTtype may occur in the stomach,

orbit, intestine, lung, thyroid, salivary gland, skin, soft tissues, bladder, kidney, andCNS. It may present as a new mass, be found on routine imaging studies, or be associated with local symptoms such as upper abdominal discomfort in gastric lymphoma. These lymphomas are localized to the organ in question in ~40% of cases and to the organ and regional lymph nodes in ~30% of patients. However, distant metastasis can occur -- particularly with transformation to diffuse large B cell lymphoma. Many patients who develop this lymphoma will have an autoimmune or inflammatory process such as Sjogren's syndrome (salivary gland MALT), Hashimoto's thyroiditis (thyroid MALT), or *Helicobacter* gastritis (gastric MALT).

Evaluation of patients with extranodal marginal zone B cell lymphoma ofMALTtype follows the pattern set forth inTable 112-11 for staging a patient with non-Hodgkin's lymphoma. In particular, patients with gastric lymphoma need to have studies performed to document the presence or absence of *H. pylori* infection. Endoscopic studies including ultrasound can help define the extent of gastric involvement. Most patients with extranodal marginal zone B cell lymphoma of MALT type have a good prognosis, with a 5-year survival of ~75%. In patients with a low IPIscore, the 5-year survival is ~90%, while it drops to ~40% in patients with a high IPI score.

TREATMENT

Extranodal marginal zone B cell lymphoma ofMALTtype is curable when localized. Local therapy such as radiation or surgery can effect cure, and this is one of the few times where surgery might be a reasonable primary therapy for a patient with non-Hodgkin's lymphoma. Patients with gastric MALT lymphomas who are infected with *H. pylori* can achieve remission in the majority of cases with eradication of the infection. These remissions can be durable. Patients who present with more extensive disease are most often treated with single-agent chemotherapy such as chlorambucil. Coexistent diffuse large B cell lymphoma must be treated with combination chemotherapy. The additional acquired mutations that mediate the histologic progression also convey *Helicobacter* independence to the growth.

Mantle Cell Lymphoma Mantle cell lymphoma makes up ~6% of all non-Hodgkin's lymphomas. Recognized as a separate entity only in the past decade, this lymphoma was previously placed in a number of other subtypes. Its existence was confirmed by the recognition that these lymphomas have a characteristic chromosomal translocation, t(11;14) between the immunoglobulin heavy chain gene on chromosome 14 and the *bcl-1* gene on chromosome 11, and regularly overexpress the BCL-1 protein. The clinical characteristics of mantle cell lymphoma are presented inTable 112-10.

The diagnosis of mantle cell lymphoma can be made accurately by an expert hematopathologist based on morphologic findings and proof that the tumor is a B cell lymphoma. As with all subtypes of lymphoma, an adequate biopsy is important. The differential diagnosis of mantle cell lymphoma includes other small-cell B cell lymphomas. In particular, mantle cell lymphoma and small lymphocytic lymphoma share a characteristic expression of CD5. Mantle cell lymphoma usually has a slightly indented nucleus.

The most common presentation of mantle cell lymphoma is with palpable

lymphadenopathy, frequently accompanied by systemic symptoms. Approximately 70% of patients will be stage IV at the time of diagnosis, with frequent bone marrow and peripheral blood involvement. Of the extranodal organs that can be involved, gastrointestinal involvement is particularly important to recognize. Patients who present with lymphomatosis polyposis in the large intestine usually have mantle cell lymphoma. The evaluation of patients with mantle cell lymphoma involves the studies presented in Table 112-11 for staging of patients with non-Hodgkin's lymphoma. Patients who present with gastrointestinal tract involvement often have Waldeyer's ring involvement, and vice versa. The 5-year survival for all patients with mantle cell lymphoma is ~25%, with only occasional patients who present with a highIPIscore surviving 5 years and ~50% of patients with a low IPI score surviving 5 years.

TREATMENT

Current therapies for mantle cell lymphoma are unsatisfactory. Patients with localized disease might be treated with combination chemotherapy followed by radiotherapy; however, these patients are exceedingly rare. For the usual presentation with disseminated disease, treatments are unsatisfactory, with the minority of patients achieving complete remission. Aggressive combination chemotherapy regimens followed by autologous or allogeneic bone marrow transplantation are frequently offered to younger patients. For the occasional elderly, asymptomatic patient, observation followed by single-agent chemotherapy might be the most practical approach. Combined used of rituximab (anti-CD20 antibody) and chemotherapy may be associated with better response rates.

Follicular Lymphoma Follicular lymphomas make up 22% of non-Hodgkin's lymphomas worldwide and at least 30% of non-Hodgkin's lymphomas diagnosed in the United States. This type of lymphoma can be diagnosed accurately on morphologic findings alone and has been the diagnosis in the majority of patients in therapeutic trials for "low-grade" lymphoma in the past. The clinical characteristics of follicular lymphoma are presented in Table 112-10.

Evaluation of an adequate biopsy by an expert hematopathologist is sufficient to make a diagnosis of follicular lymphoma. The tumor is composed of small cleaved and large cells in varying proportions organized in a follicular pattern of growth (see Plate V-30). Confirmation of B cell immunophenotype and the existence of the t(14;18) and abnormal expression of BCL-2 protein are confirmatory. The major differential diagnosis is between lymphoma and reactive follicular hyperplasia. The coexistence of diffuse large B cell lymphoma must be considered. Patients with follicular lymphoma are often subclassified into those with predominantly small cells, those with a mixture of small and large cells, and those with predominantly large cells. While this distinction cannot be made simply or very accurately, these subdivisions do have prognostic significance. Patients with follicular lymphoma with predominantly large cells have a higher proliferative fraction, progress more rapidly, and have a shorter overall survival with simple chemotherapy regimens.

The most common presentation for follicular lymphoma is with new, painless lymphadenopathy. Multiple sites of lymphoid involvement are typical, and unusual sites such as epitrochlear nodes are sometimes seen. However, essentially any organ can be

involved, and extranodal presentations do occur. Most patients do not have fevers, sweats, or weight loss, and anIPIscore of 0 or 1 is found in ~50% of patients. Fewer than 10% of patients have a high (i.e., 4 or 5) IPI score. The staging evaluation for patients with follicular lymphoma should include the studies included inTable 112-11 for the staging of patients with non-Hodgkin's lymphoma.

TREATMENT

Follicular lymphoma is one of the malignancies most responsive to chemotherapy and radiotherapy. In addition, as many as 25% of the patients undergo spontaneous regression -- usually transient -- when followed without therapy. In an asymptomatic patient, no initial treatment and watchful waiting can be an appropriate management strategy and is particularly likely to be adopted for older patients. For patients who do require treatment, single-agent chlorambucil or cyclophosphamide or combination chemotherapy withCVP orCHOP are most frequently used. With adequate treatment, between 50 and 75% of patients will achieve a complete remission. While most patients relapse (median response duration is about 2 years), at least 20% of complete responders will remain in remission for >10 years. For the rare patient with localized follicular lymphoma, involved field radiotherapy produces an excellent treatment result.

A number of new therapies have been shown to be active in the treatment of patients with follicular lymphoma. These include new cytotoxic agents such as fludarabine, interferona, monoclonal antibodies with or without radionuclides, and lymphoma vaccines. In patients treated with a doxorubicin-containing combination chemotherapy regimen, interferona given to patients in complete remission seems to prolong survival. The monoclonal antibody rituximab can cause objective responses in 35 to 50% of patients with relapsed follicular lymphoma, and radiolabeled antibodies appear to have response rates well in excess of 50%. Trials with tumor vaccines have been encouraging. Both autologous and allogeneic hematopoietic stem cell transplantation yield high complete response rates in patients with relapsed follicular lymphoma, and long-term remissions can occur.

Patients with follicular lymphoma with a predominance of large cells have a shorter survival when treated with single-agent chemotherapy but seem to benefit from receiving an anthracycline-containing combination chemotherapy regimen. When their disease is treated aggressively, the overall survival for such patients is no lower than for patients with other follicular lymphomas, and the failure-free survival is superior.

Patients with follicular lymphoma have a high rate of histologic transformation to diffuse large B cell lymphoma (~7% per year). This is recognized ~40% of the time during the course of the illness by repeat biopsy and is present in almost all patients at autopsy. This transformation is usually heralded by rapid growths of lymph nodes -- often localized -- and the development of systemic symptoms such as fevers, sweats, and weight loss. Although these patients have a poor prognosis, aggressive combination chemotherapy regimens can sometimes cause a complete remission in the diffuse large B cell lymphoma, usually leaving the patient with persisting follicular lymphoma.

Diffuse Large B Cell Lymphoma Diffuse large B cell lymphoma is the most common type of non-Hodgkin's lymphoma, representing approximately one-third of all cases.

This lymphoma makes up the majority of cases in previous clinical trials of "aggressive" or "intermediate-grade" lymphoma. The clinical characteristics of diffuse large B cell lymphoma are presented in Table 112-10.

The diagnosis of diffuse large B cell lymphoma can be made accurately by an expert hematopathologist when review of an adequate biopsy and proof of B cell immunophenotype are available (seePlate V-22). Cytogenetic and molecular genetic studies are not necessary for diagnosis, but some evidence has accumulated that patients who overexpress the BCL-2 protein might be more likely to relapse than others. Patients with prominent mediastinal involvement are sometimes diagnosed as a separate subgroup having primary mediastinal diffuse large B cell lymphoma. This latter group of patients has a younger median age (i.e., 37 years) and a female predominance (66%). Subtypes of diffuse large B cell lymphoma, including those with an immunoblastic subtype and tumors with extensive fibrosis, are recognized by pathologists but do not appear to have important, independent prognostic significance.

Diffuse large B cell lymphoma can present as either primary lymph node disease or at extranodal sites. More than 50% of patients will have some site of extranodal involvement at diagnosis, with the most common sites being the gastrointestinal tract and bone marrow, each being involved in 15 to 20% of patients. Essentially any organ can be involved, making a diagnostic biopsy imperative. For example, diffuse large B cell lymphoma of the pancreas has a much better prognosis than pancreatic carcinoma but would be missed without biopsy. Primary diffuse large B cell lymphoma of the brain is being diagnosed with increasing frequency.

The initial evaluation of patients with diffuse large B cell lymphoma involves the studies presented inTable 112-11 for staging of patients with non-Hodgkin's lymphoma. After a careful staging evaluation, ~50% of patients will be found to have stage I or II disease and ~50% will have widely disseminated lymphoma. Bone marrow biopsy shows involvement by lymphoma in about 15% of cases, with marrow involvement by small cells more frequent than with large cells.

TREATMENT

The initial treatment of all patients with diffuse large B cell lymphoma should be with a combination chemotherapy regimen. The most popular regimen in the United States isCHOP, although a variety of other anthracycline-containing combination chemotherapy regimens appear to be equally efficacious. Patients with stage I or nonbulky stage II can be effectively treated with three to four cycles of combination chemotherapy followed by involved field radiotherapy. The results are at least equal and probably superior to six to eight cycles of combination therapy, and cure rates of 60 to 70% in stage II disease and 80 to 90% in stage I disease can be expected.

For patients with bulky stage II, stage III, or stage IV, six to eight cycles of combination chemotherapy regimen such asCHOP are usually administered. A frequent approach would be to administer four cycles of therapy and then reevaluate. If the patient has achieved a complete remission after four cycles, two more cycles of treatment might be given and then therapy discontinued. Using this approach, ~60 to 70% of patients can be expected to achieve a complete remission, and 50 to 70% of complete responders

will be cured. The chances for a favorable response to treatment are predicted by theIPI. In fact, the IPI was developed specifically to predict outcome in patients with diffuse large-cell lymphoma. For the 35% of patients with a low IPI score of 0 to 1, the 5-year survival is >70%, while for the 20% of patients with a high IPI score of 4 to 5, the 5-year survival is ~20%. A number of other factors, including molecular features of the tumor, levels of circulating cytokines and soluble receptors, and other surrogate markers, have been shown to influence prognosis. However, they have not been validated as rigorously as the IPI and have not been uniformly applied clinically.

Because a large number of patients with diffuse large B cell lymphoma are either initially refractory to therapy or relapse after apparently effective chemotherapy,>50% of patients will be candidates for salvage treatment at some point. Alternative combination chemotherapy regimens can induce complete remission in as many as 50% of these patients, but long-term disease-free survival is seen in £10%. Autologous bone marrow transplantation has been shown to be superior to salvage chemotherapy at usual doses and leads to long-term disease-free survival in ~40% of patients whose lymphomas remain chemotherapy-sensitive after relapse.

Burkitt's Lymphoma/Leukemia Burkitt's lymphoma/leukemia is a rare disease in adults in the United States, making up <1% of non-Hodgkin's lymphomas, but it makes up ~30% of childhood non-Hodgkin's lymphoma. Burkitt's leukemia, or L3 ALL, makes up a small proportion of childhood and adult acute leukemias. The clinical features of Burkitt's lymphoma occurring in adults are presented inTable 112-10.

Burkitt's lymphoma can be diagnosed morphologically by an expert hematopathologist with a high degree of accuracy. The cells are homogeneous in size and shape (seePlate V-4). Demonstration of a very high proliferative fraction and the presence of the t(8;14) or one of its variants, t(2;8) (*c-myc* and the l light chain gene) or t(8;22) (*c-myc* and the k light chain gene), can be confirmatory. Burkitt's cell leukemia is recognized by the typical medium-sized cells with round nuclei, multiple nucleoli, and basophilic cytoplasm with cytoplasmic vacuoles. Demonstration of a B cell immunophenotype and one of the above-noted cytogenetic abnormalities is confirmatory.

The three distinct clinical forms of Burkitt's lymphoma that are recognized are endemic, sporadic, and immunodeficiency-associated. Endemic and sporadic Burkitt's lymphomas occur frequently in children in Africa, and the sporadic form in western countries. Immunodeficiency-associated Burkitt's lymphoma is seen in patients with HIV infection.

Pathologists sometimes have difficulty distinguishing between Burkitt's lymphoma and diffuse large B cell lymphoma. In the past, a separate subgroup of non-Hodgkin's lymphoma intermediate between the two was recognized. When tested, this subgroup could not be diagnosed accurately. Distinction between the two major types of B cell aggressive non-Hodgkin's lymphoma can sometimes be made based on the extremely high proliferative fraction seen in patients with Burkitt's lymphoma (i.e., essentially 100% of tumor cells are in cycle) caused by *c-myc* deregulation.

Most patients in the United States with Burkitt's lymphoma present with peripheral lymphadenopathy or an intraabdominal mass. The disease is typically rapidly progressive and has a propensity to metastasize to theCNS. Initial evaluation should

always include an examination of cerebral spinal fluid to rule out metastasis in addition to the other staging evaluations noted inTable 112-11. Once the diagnosis of Burkitt's lymphoma is suspected, a diagnosis must be made promptly and staging evaluation must be accomplished expeditiously. This is the most rapidly progressive human tumor, and any delay in initiating therapy can adversely affect the patient's prognosis.

TREATMENT

Treatment of Burkitt's lymphoma in both children and adults involves the use of intensive combination chemotherapy regimens incorporating administered high doses of cyclophosphamide. Prophylactic therapy to theCNS is mandatory and incorporated in all modern regimens. Burkitt's lymphoma was one of the first cancers shown to be curable by chemotherapy. Today, cure can be expected in 70% of both children and young adults when effective therapy is administered precisely. Salvage therapy has been generally ineffective in patients failing the initial treatment, emphasizing the importance of the initial treatment approach.

Other B Cell Lymphoid Malignancies *B-cell prolymphocytic leukemia* involves blood and marrow infiltration by large lymphocytes with prominent nucleoli. Patients typically have a high white cell count, splenomegaly, and minimal lymphadenopathy. The chances for a complete response to therapy are poor.

Hairy cell leukemia is a rare disease that presents predominantly in older males. Typical presentation involves pancytopenia, although occasional patients will have a leukemic presentation. Splenomegaly is usual. The malignant cells appear to have "hairy" projections on light and electron microscopy and show a characteristic staining pattern with tartrate-resistant acid phosphatase. Bone marrow is typically not able to be aspirated, and biopsy shows a pattern of fibrosis with diffuse infiltration by the malignant cells. Patients with this disorder are prone to unusual infections including infection by *Mycobacterium avium intracellulare*, and vasculitic syndromes have been described. Hairy cell leukemia is responsive to chemotherapy with interferona, pentostatin, or cladribine, with the latter being the usually preferred treatment. Clinical complete remissions with cladribine occur in the majority of patients, and long-term disease-free survival is frequent.

Splenic marginal zone lymphoma involves infiltration of the splenic white pulp by small, monoclonal B lymphocytes. This is a rare disorder that can present as leukemia as well as lymphoma. Definitive diagnosis is often made at splenectomy, which is also an effective therapy. This is an extremely indolent disorder, but when chemotherapy is required, the most usual treatment has been chlorambucil.

Lymphoplasmacytic lymphoma is the tissue manifestation of Waldenstrom's macroglobulinemia (Chap. 113). This type of lymphoma has been associated with chronic hepatitis C virus infection, and an etiologic association has been proposed. Patients typically present with lymphadenopathy, splenomegaly, bone marrow involvement, and occasionally peripheral blood involvement. The tumor cells do not express CD5. Patients often have a monoclonal IgM protein, high levels of which can dominate the clinical picture with the symptoms of hyperviscosity. Treatment of lymphoplasmacytic lymphoma can be aimed primarily at reducing the abnormal protein,

if present, but will usually also involve chemotherapy. Chlorambucil, fludarabine, and cladribine have been utilized. The median 5-year survival for patients with this disorder is ~60%.

Nodal marginal zone lymphoma, also known as *monocytoid B cell lymphoma*, represents ~1% of non-Hodgkin's lymphomas. This lymphoma has a slight female predominance and presents with disseminated disease (i.e., stage III or IV) in 75% of patients. Approximately one-third of patients have bone marrow involvement, and a leukemic presentation occasionally occurs. The staging evaluation and therapy should use the same approach as used for patients with follicular lymphoma. Approximately 60% of the patients with nodal marginal zone lymphoma will survive 5 years after diagnosis.

PRECURSOR CELL T CELL MALIGNANCIES

Precursor T Cell Lymphoblastic Leukemia/Lymphoma Precursor T cell malignancies can present either as ALL or as an aggressive lymphoma. These malignancies are more common in children and young adults, with males more frequently affected than females.

Precursor T cell ALL can present with bone marrow failure, although the severity of anemia, neutropenia, and thrombocytopenia is often less than in precursor B cell ALL. These patients sometimes have very high white cell counts, a mediastinal mass, lymphadenopathy, and hepatosplenomegaly. Precursor T cell lymphoblastic lymphoma is most often found in young men presenting with a large mediastinal mass and pleural effusions. Both presentations have a propensity to metastasize to theCNS, and CNS involvement is often present at diagnosis.

TREATMENT

Children with precursor T cell ALL seem to benefit from very intensive remission induction and consolidation regimens. The majority of patients treated in this manner can be cured. Older children and young adults with precursor T cell lymphoblastic lymphoma are also often treated with "leukemia-like" regimens. Patients who present with localized disease have an excellent prognosis. However, advanced age is an adverse prognostic factor. Adults with precursor T cell lymphoblastic lymphoma who present with highLDHlevels or bone marrow orCNSinvolvement are often offered bone marrow transplantation as part of their primary therapy.

MATURE (PERIPHERAL) T CELL DISORDERS

Mycosis Fungoides Mycosis fungoides is also known as *cutaneous T cell lymphoma*. This lymphoma is more often seen by dermatologists than internists. The median age of onset is in the mid-fifties, and the disease is more common in males and in blacks.

Mycosis fungoides is an indolent lymphoma with patients often having several years of eczematous or dermatitic skin lesions before the diagnosis is finally established. The skin lesions progress from patch stage to plaque stage to cutaneous tumors. Early in the disease, biopsies are often difficult to interpret, and the diagnosis may only become

apparent by observing the patient over time. In advanced stages, the lymphoma can metastasize to lymph nodes and visceral organs. A particular syndrome in patients with this lymphoma involves erythroderma and circulating tumor cells. This is known as Sezary's syndrome.

Rare patients with localized early stage mycosis fungoides can be cured with radiotherapy, often total-skin electron beam irradiation. More advanced disease has been treated with topical glucocorticoids, topical nitrogen mustard, phototherapy, psoralen with ultraviolet A (PUVA), electron beam radiation, interferon, and systemic cytotoxic therapy. Unfortunately, these treatments are palliative.

Adult T Cell Lymphoma/Leukemia Adult T cell lymphoma/leukemia is one manifestation of infection by theHTLV-I retrovirus. Patients can be infected through transplacental transmission, blood transfusion, and by sexual transmission of the virus. Patients who acquire the virus from their mother through breast milk are most likely to develop lymphoma, but the risk is still only 2.5% and the latency averages 55 years. Tropical spastic paraparesis, another manifestation of HTLV-I infection (Chap. 191), occurs after a shorter latency (1 to 3 years) and is most common in people who acquire the virus during adulthood from transfusion or sex.

The diagnosis of adult T cell lymphoma/leukemia is made when an expert hematopathologist recognizes the typical morphologic picture, a T cell immunophenotype (i.e., CD4 positive) of malignant cells has been demonstrated, and the existence of antibodies to HTLV-I is proven. Examination of the peripheral blood will usually reveal characteristic, pleomorphic abnormal CD4-positive cells with indented nuclei, which have been called "flower" cells (seePlate V-40).

A subset of patients have a smoldering clinical course and long survival, but most patients present with an aggressive disease manifested by lymphadenopathy, hepatosplenomegaly, skin infiltration, hypercalcemia, lytic bone lesions, and elevatedLDHlevels. The skin lesions can be papules, plaques, tumors, and ulcerations. Bone marrow involvement is not usually extensive, and anemia and thrombocytopenia are not usually prominent. Although treatment by combination chemotherapy regimens can result in objective responses, true complete remissions are unusual, and the median survival of patients is about 7 months.

Anaplastic Large T/Null Cell Lymphoma Anaplastic large T/null cell lymphoma was previously usually diagnosed as undifferentiated carcinoma or malignant histiocytosis. Discovery of the CD30, or Ki-1, antigen and the recognition that some patients with previously unclassified malignancies displayed this antigen led to the identification of a new type of lymphoma. Subsequently, discovery of the t(2;5) and the resultant frequent overexpression of the anaplastic lymphoma kinase (alk) protein confirmed the existence of this entity. This lymphoma accounts for ~2% of all non-Hodgkin's lymphomas. The clinical characteristics of patients with anaplastic large T/null cell lymphoma are presented inTable 112-10.

The diagnosis of anaplastic large T/null cell lymphoma is made when an expert hematopathologist recognizes the typical morphologic picture and a T cell or null cell immunophenotype is demonstrated along with CD30 positivity. Documentation of the

t(2;5) and/or overexpression of alk protein confirm the diagnosis. Some diffuse large B cell lymphomas can also have an anaplastic appearance but have the same clinical course or response to therapy as other diffuse large B cell lymphomas.

Patients with anaplastic large T/cell null cell lymphoma are typically young (median age, 33 years) and male (~70%). Some 50% of patients present in stage I/II, and the remainder with more extensive disease. Systemic symptoms and elevatedLDHlevels are seen in about one-half of patients. Bone marrow and the gastrointestinal tract are rarely involved, but skin involvement is frequent. Some patients with disease confined to the skin have a different and more indolent disorder that has been termed *cutaneous anaplastic large T/null cell lymphoma* and might be related to lymphomatoid papulosis.

TREATMENT

Treatment regimens appropriate for other aggressive lymphomas, such as diffuse large B cell lymphoma, should be utilized in patients with anaplastic large T/null cell lymphoma. Surprisingly, given the anaplastic appearance, this disorder has the best survival rate of any aggressive lymphoma. The 5-year survival is>75%. While traditional prognostic factors such as the IPIpredict treatment outcome, overexpression of the alk protein is an important prognostic factor, with patients overexpressing this protein having a superior treatment outcome.

Peripheral T Cell Lymphoma The peripheral T cell lymphomas make up a heterogenous morphologic group of aggressive neoplasms that share a mature T cell immunophenotype. They represent ~7% of all cases of non-Hodgkin's lymphoma. A number of distinct clinical syndromes are included in this group of disorders. The clinical characteristics of patients with peripheral T cell lymphoma are presented inTable 112-10.

The diagnosis of peripheral T cell lymphoma, or any of its specific subtypes, requires an expert hematopathologist, an adequate biopsy, and immunophenotyping. Most peripheral T cell lymphomas are CD4-positive, but a few will be CD8-positive, both CD4- and CD8-positive, or have anNK-cell immunophenotype. No characteristic genetic abnormalities have yet been identified, but translocations involving the T cell antigen receptor genes on chromosomes 7 or 14 may be detected. The differential diagnosis of patients suspected of having peripheral T cell lymphoma includes reactive T cell infiltrative processes. In some cases, demonstration of a monoclonal T cell population using T cell receptor gene rearrangement studies will be required to make a diagnosis.

The initial evaluation of a patient with a peripheral T cell lymphoma should include the studies inTable 112-11 for staging patients with non-Hodgkin's lymphoma. Unfortunately, patients with peripheral T cell lymphoma usually present with adverse prognostic factors, with >80% of patients having anIPIscore ³2 and >30% having an IPI score ³4. As this would predict, peripheral T cell lymphomas are associated with a poor outcome, and only 25% of the patients survive 5 years after diagnosis. Treatment regimens are the same as those used for diffuse large B cell lymphoma, but patients with peripheral T cell lymphoma have a poorer response to treatment. Because of this poor treatment outcome, hematopoietic stem cell transplantation is often considered early in the care of young patients.

A number of specific clinical syndromes are seen in the peripheral T cell lymphomas. *Angioimmunoblastic T cell lymphoma* is one of the more common subtypes, making up ~20% of T cell lymphomas. These patients typically present with generalized lymphadenopathy, fever, weight loss, skin rash, and polyclonal hypergammaglobunenemia. In some cases, it is difficult to separate patients with a reactive disorder from those with true lymphoma.

Extranodal T/NK cell lymphoma of nasal type has also been called *angiocentric lymphoma* and was previously termed *lethal midline granuloma*. This disorder is more frequent in Asia and South America than in the United States and Europe. Although most frequent in the upper airway, it can involve other organs. The course is aggressive, and patients frequently have the hemophagocytic syndrome. When marrow and blood involvement occur, distinction between this disease and leukemia might be difficult. Some patients will respond to aggressive combination chemotherapy regimens, but the overall outlook is poor.

Enteropathy-type intestinal T cell lymphoma is a rare disorder that occurs in patients with untreated gluten-sensitive enteropathy. Patients are frequently wasted and sometimes present with intestinal perforation. The prognosis is poor. *Hepatosplenicgd T cell lymphoma* is a systemic illness that presents with sinusoidal infiltration of the liver, spleen, and bone marrow by malignant T cells. Tumor masses generally do not occur. The disease is associated with systemic symptoms and is often difficult to diagnosis. Treatment outcome is poor. *Subcutaneous panniculitis-like T cell lymphoma* is a rare disorder that is often confused with panniculitis. Patients present with multiple subcutaneous nodules, which progress and can ulcerate. Hemophagocytic syndrome is common. Response to therapy is poor. The development of the hemophagocytic syndrome (profound anemia, ingestion of erythrocytes by monocytes and macrophages) in the course of any peripheral T cell lymphoma is generally associated with a fatal outcome.

HODGKIN'S DISEASE

Nodular Lymphocyte-Predominant Hodgkin's Disease Nodular lymphocyte predominant Hodgkin's disease is now recognized as an entity distinct from classic Hodgkin's disease. Previous classification systems recognized that biopsies from a subset of patients diagnosed as having Hodgkin's disease contained a predominance of small lymphocytes and rare Sternberg-Reed cells. In recent years, it was recognized that a subset of these patients had a nodular growth pattern and a clinical course that varied from that of patients with classic Hodgkin's disease. This is an unusual clinical entity and represents <5% of cases of Hodgkin's disease.

Nodular lymphocyte-predominant Hodgkin's disease has a number of characteristics that suggest its relationship to non-Hodgkin's lymphoma. These include a clonal proliferation of B cells and a distinctive immunophenotype; tumor cells express J chain and display CD45 and epithelial membrane antigen (ema) and do not express two markers normally found on Sternberg-Reed cells, CD30 and CD15. This lymphoma tends to have a chronic, relapsing course and sometimes transforms to diffuse large B cell lymphoma.

The treatment of patients with nodular lymphocyte-predominant Hodgkin's disease is controversial. Some clinicians favor no treatment and merely close follow-up. In the United States, most physicians will treat localized disease with radiotherapy and disseminated disease with regimens utilized for patients with classic Hodgkin's disease. Regardless of the therapy utilized, most series report a long-term survival of>80%.

Classical Hodgkin's Disease Hodgkin's disease occurs in ~8000 patients in the United States each year, and the disease does not appear to be increasing in frequency. Most patients present with palpable lymphadenopathy that is nontender; in most patients, these lymph nodes are in the neck, supraclavicular area, and axilla. More than half the patients will have mediastinal adenopathy at diagnosis, and this is sometimes the initial manifestation. Subdiaphragmatic presentation of Hodgkin's disease is unusual and more common in older males. Approximately one-third of patients present with fevers, night sweats, and/or weight loss -- B symptoms in the Ann Arbor staging classification (Table 112-8). Occasionally, Hodgkin's disease can present as a fever of unknown origin. This is more common in older patients who are found to have mixed-cellularity Hodgkin's disease in an abdominal site. Rarely, the fevers persist for days to weeks, followed by afebrile intervals and then recurrence of the fever. This pattern is known as *Pel-Epstein fever*. Hodgkin's disease can occasionally present with unusual manifestations. These include severe and unexplained itching, cutaneous disorders such as erythema nodosum and ichthyosiform atrophy, paraneoplastic cerebellar degeneration and other distant effects on theCNS, nephrotic syndrome, immune hemolytic anemia and thrombocytopenia, hypercalcemia, and pain in lymph nodes on alcohol ingestion.

The diagnosis of Hodgkin's disease is established by review of an adequate biopsy specimen by an expert hematopathologist. In the United States, most patients would be classified as having nodular sclerosing Hodgkin's disease, with a minority of patients having mixed-cellularity Hodgkin's disease. Lymphocyte-predominant and lymphocyte-depleted Hodgkin's disease are rare. Mixed-cellularity Hodgkin's disease or lymphocyte-depletion Hodgkin's disease are seen more frequently in patients infected by HIV (see Plate V-18). The differential diagnosis of a lymph node biopsy suspicious for Hodgkin's disease includes inflammatory processes, mononucleosis, non-Hodgkin's lymphoma, diphenylhydantoin-induced lymphadenopathy, and nonlymphomatous malignancies.

The staging evaluation for a patient with Hodgkin's disease would typically include a careful history and physical examination; complete blood count; erythrocyte sedimentation rate; serum chemistry studies includingLDH; chest radiograph;CT scan of the chest, abdomen, and pelvis; and bone marrow biopsy. Many patients would also have a gallium scan. If radiologic expertise is available, a bipedal lymphangiogram can be helpful. Gallium scans are most useful at the completion of therapy to document remission. Staging laparotomies were once popular for most patients with Hodgkin's disease but are now done rarely because of an increased reliance on systemic rather than local therapy.

TREATMENT

Patients with localized Hodgkin's disease are cured >90% of the time. In patients with good prognostic factors, extended field radiotherapy has a high cure rate. Increasingly, patients with all stages of Hodgkin's disease are treated initially with chemotherapy. Patients with localized or good-prognosis disease receive a brief course of chemotherapy followed by radiotherapy to sites of node involvement. Patients with more extensive disease or those with B symptoms receive a complete course of chemotherapy. The most popular chemotherapy regimens used in the treatment of Hodgkin's disease include doxorubicin, bleomycin, vinblastine, and dacarbazine (ABVD) and mechlorethamine, vincristine, procarbazine, and prednisone (MOPP), or combinations of the drugs in these two regimens. Today, most patients in the United States receive ABVD. Long-term disease-free survival in patients with advanced disease can be achieved in >75% of patients who lack systemic symptoms and in 50 to 70% of patients with systemic symptoms.

Patients who relapse after primary therapy of Hodgkin's disease can frequently still be cured. Patients who relapse after initial treatment only with radiotherapy have an excellent outcome when treated with chemotherapy. Patients who relapse after an effective chemotherapy regimen are usually not curable with subsequent chemotherapy administered at standard doses. However, patients with a long initial remission can be an exception to this rule. Autologous bone marrow transplantation can cure half of patients who fail effective chemotherapy regimens.

Because of the very high cure rate in patients with Hodgkin's disease, long-term complications have become a major focus for clinical research. In fact, in some series of patients with early-stage disease, more patients died from late complications of therapy than from Hodgkin's disease itself. This is particularly true in patients with localized disease. The most serious late side effects include second malignancies and cardiac injury. Patients are at risk for the development of acute leukemia in the first 10 years after treatment with combination chemotherapy regimens that contain alkylating agents. The risk for development of acute leukemia appears to be greater after MOPP-like regimens than with ABVD. The development of carcinomas as a complication of treatment for Hodgkin's disease has become a major problem. These tumors usually occur[3]10 years after treatment and are associated more with radiotherapy than with chemotherapy. For this reason, young women treated with thoracic radiotherapy for Hodgkin's disease should institute screening mammograms 5 to 10 years after treatment, and all patients who receive thoracic radiotherapy for Hodgkin's disease should be discouraged from smoking. Thoracic radiation also accelerates coronary artery disease, and patients should be encouraged to minimize risk factors for coronary artery disease such as smoking and elevated cholesterol levels.

A number of other late side effects from the treatment of Hodgkin's disease are well known. Patients who receive thoracic radiotherapy are at very high risk for the eventual development of hypothyroidism and should be observed for this complication; intermittent measurement of thyrotropin should be made to identify the condition before it becomes symptomatic. Lhermitte's syndrome occurs in ~15% of patients who receive thoracic radiotherapy. This syndrome is manifested by an "electric shock" sensation into the lower extremities on flexion of the neck. Infertility is a concern for all patients undergoing treatment for Hodgkin's disease. In both women and men, the risk of permanent infertility is age-related, with younger patients more likely to recover fertility.

In addition, treatment withABVDrather thanMOPPincreases the chances to retain fertility.

LYMPHOMA-LIKE DISORDERS

The most common condition that pathologists and clinicians might confuse with lymphoma is reactive, atypical lymphoid hyperplasia. Patients might have localized or disseminated lymphadenopathy and might have the systemic symptoms characteristic of lymphoma. Underlying causes include a drug reaction to diphenylhydantoin or carbamezepine. Immune disorders such as rheumatoid arthritis and lupus erythematosus, viral infections such as cytomegalovirus andEBV, and bacterial infections such as cat-scratch disease may cause adenopathy (Chap. 63). In the absence of a definitive diagnosis after initial biopsy, continued follow-up, further testing, and repeated biopsies, if necessary, are the appropriate approach rather than instituting therapy.

Specific conditions that can be confused with lymphoma include *Castleman's disease*, which can present with localized or disseminated lymphadenopathy; some patients have systemic symptoms. The disseminated form is often accompanied by anemia and polyclonal hypergammaglobulinemia, and the condition seems to be related to an overproduction of interleukin 6, possibly produced by human herpesvirus 8. Patients with localized disease can be treated effectively with local therapy, while the initial treatment for patients with disseminated disease is usually with systemic glucocorticoids.

Sinus histiocytosis with massive lymphadenopathy (*Rosai-Dorfman's disease*) usually presents with bulky lymphadenopathy in children or young adults. The disease is usually nonprogressive and self-limited, but patients can manifest autoimmune hemolytic anemia.

Lymphomatoid papulosis is a cutaneous lymphoproliferative disorder that is often confused with anaplastic large-cell lymphoma involving the skin. The cells of lymphomatoid papulosis are similar to those seen in lymphoma and stain for CD30, and T cell receptor gene rearrangements are sometimes seen. However, the condition is characterized by waxing and waning skin lesions that usually heal, leaving small scars. In the absence of effective communication between the clinician and the pathologist regarding the clinical course in the patient, this disease will be misdiagnosed. Since the clinical picture is usually benign, misdiagnosis is a serious mistake.

ACKNOWLEDGEMENT

Dr. Arnold Freedman and Dr. Lee Nadler contributed this chapter to the 14th edition, and some elements of that chapter were retained here.

(Bibliography omitted in Palm version)

Back to Table of Contents

113. PLASMA CELL DISORDERS - *Dan L. Longo*

GENERAL PRINCIPLES

The *plasma cell disorders* are monoclonal neoplasms related to each other by virtue of their development from common progenitors in the B lymphocyte lineage. Multiple myeloma, Waldenstrom's macroglobulinemia, primary amyloidosis (Chap. 319), and the heavy chain diseases comprise this group and may be designated by a variety of synonyms such as *monoclonal gammopathies*, *paraproteinemias*, *plasma cell dyscrasias*, and *dysproteinemias*. Mature B lymphocytes destined to produce IgG bear surface immunoglobulin molecules of both M and G heavy chain isotypes with both isotypes having identical idiotypes (variable regions). Under normal circumstances, maturation to antibody-secreting plasma cells is stimulated by exposure to the antigen for which the surface immunoglobulin is specific; however, in the plasma cell disorders the control over this process is lost. The clinical manifestations of all the plasma cell disorders relate to the expansion of the neoplastic cells, to the secretion of cell products (immunoglobulin molecules or subunits, lymphokines), and to some extent to the host's response to the tumor.**Normal development of B lymphocytes is discussed in Chap. 305.*

There are three categories of structural variation among immunoglobulin molecules that form antigenic determinants, and these are used to classify immunoglobulins (Chap. 305). *Isotypes* are those determinants that distinguish among the main classes of antibodies of a given species and are the same in all normal individuals of that species. Therefore, isotypic determinants are, by definition, recognized by antibodies from a distinct species (heterologous sera) but not by antibodies from the same species (homologous sera). There are five heavy chain isotypes (M, G, A, D, E) and two light chain isotypes (k, l). *Allotypes* are distinct determinants that reflect regular small differences between individuals of the same species in the amino acid sequences of otherwise similar immunoglobulins. These differences are determined by allelic genes; by definition, they are detected by antibodies made in the same species. *Idiotypes* are the third category of antigenic determinants. They are unique to the molecules produced by a given clone of antibody-producing cells. Idiotypes are formed by the unique structure of the antigen-binding portion of the molecule.

Antibody molecules (Fig. 305-8) are composed of two heavy chains (mol wt ~50,000) and two light chains (mol wt ~25,000). Each chain has a constant portion (limited amino acid sequence variability) and a variable region (extensive sequence variability). The light and heavy chains are linked by disulfide bonds and are aligned so that their variable regions are adjacent to one another. This variable region forms the antigen recognition site of the antibody molecule; its unique structural features form a particular set of determinants, or idiotypes, that are reliable markers for a particular clone of cells because each antibody is formed and secreted by a single clone. Each chain is specified by distinct genes, synthesized separately, and assembled into an intact antibody molecule after translation (Fig. 113-1). Because of the mechanics of the gene rearrangements necessary to specify the immunoglobulin variable regions (VDJ joining for the heavy chain, VJ joining for the light chain), a particular clone rearranges only one of the two chromosomes to produce an immunoglobulin molecule of only one light chain isotype and only one allotype (allelic exclusion). After exposure to antigen, the variable

region may become associated with a new heavy chain isotype (class switch). Each clone of cells performs these sequential gene arrangements in a unique way. This results in each clone producing a unique immunoglobulin molecule. In most cells, light chains are synthesized in slight excess, are secreted as free light chains by plasma cells, and are cleared by the kidney, but <10 mg of such light chains is excreted per day.

Electrophoretic analysis of components of the serum proteins permits determination of the amount of immunoglobulin in the serum (Fig. 113-2). The variety of immunoglobulins move heterogeneously in an electric field and form a broad peak in the gamma region. The gamma globulin region of the electrophoretic pattern is usually increased in the sera of patients and animals with plasma cell tumors. There is a sharp spike in this region called an *M component* (M for monoclonal). Less commonly, the M component may appear in the $beta_2$ or $alpha_2$globulin region. The antibody must be present at a concentration of at least 5 g/L (0.5 g/dL) to be detectable by this method. This corresponds to approximately 10_9cells producing the antibody. Confirmation that such an M component is truly monoclonal relies on the use of immunoelectrophoresis that shows a single light and heavy chain type. Hence immunoelectrophoresis and electrophoresis provide qualitative and quantitative assessment of the M component, respectively. Once the presence of an M component has been confirmed, electrophoresis provides the more practical information for managing patients with monoclonal gammopathies. In a given patient, the amount of M component in the serum is a reliable measure of the tumor burden. This makes the M component an excellent tumor marker, yet it is not specific enough to be used to screen asymptomatic patients. In addition to the plasma cell disorders, M components may be detected in other lymphoid neoplasms such as chronic lymphocytic leukemia and lymphomas of B or T cell origin; nonlymphoid neoplasms such as chronic myeloid leukemia, breast cancer, and colon cancer; a variety of nonneoplastic conditions such as cirrhosis, sarcoidosis, parasitic diseases, Gaucher's disease, and pyoderma gangrenosum; and a number of autoimmune conditions, including rheumatoid arthritis, myasthenia gravis, and cold agglutinin disease. A very rare skin disease known as lichen myxedematosus or papular mucinosis is associated with a monoclonal gammopathy. Highly cationic IgG is deposited in the dermis of patients with this disease. This organ specificity may reflect the specificity of the antibody for some antigenic component of the dermis.

The nature of the M component is variable in plasma cell disorders. It may be an intact antibody molecule of any heavy chain subclass, or it may be an altered antibody or fragment. Isolated light or heavy chains may be produced. In some plasma cell tumors such as extramedullary or solitary bone plasmacytomas, less than a third of patients will have an M component. In about 20% of myelomas, only light chains are produced and in most cases are secreted in the urine as Bence Jones proteins. The frequency of myelomas of a particular heavy chain class is roughly proportional to the serum concentration, and therefore IgG myelomas are more common than IgA and IgD myelomas.

MULTIPLE MYELOMA

Definition Multiple myeloma represents a malignant proliferation of plasma cells derived from a single clone. The terms *multiple myeloma* and *myeloma* may be used

interchangeably. The tumor, its products, and the host response to it result in a number of organ dysfunctions and symptoms of bone pain or fracture, renal failure, susceptibility to infection, anemia, hypercalcemia, and occasionally clotting abnormalities, neurologic symptoms, and vascular manifestations of hyperviscosity.

Etiology The cause of myeloma is not known. Myeloma occurred with increased frequency in those exposed to the radiation of nuclear warheads in World War II after a 20-year latency. A variety of chromosomal alterations have been found in patients with myeloma; 13q14 deletions, 17p13 deletions, and 11q abnormalities predominate. The most common translocation is t(11;14)(q13;q32), and evidence is strong that errors in switch recombination -- the genetic mechanism to change antibody heavy chain isotype -- participate in the transformation pathway. Overexpression of *myc* or *ras* genes has been noted in some cases. Mutations in p53 and Rb-1 have also been described, but no common molecular pathogenesis has yet emerged.

Myeloma has been seen more commonly than expected among farmers, wood workers, leather workers, and those exposed to petroleum products. The neoplastic event in myeloma may involve cells earlier in B cell differentiation than the plasma cell. Circulating B cells bearing surface immunoglobulin that share the idiotype of the M component are present in myeloma patients. Interleukin (IL) 6 may play a role in driving myeloma cell proliferation; a large fraction of myeloma cells exposed to IL-6 in vitro respond by proliferating. The IL-6 dependency of myeloma is controversial. Infection of marrow macrophages with human herpesvirus 8 has been noted in some cases leading to the hypothesis that viral IL-6 may contribute to the pathogenesis. This notion is also debated. It remains difficult to distinguish benign from malignant plasma cells on the basis of morphologic criteria in all but a few cases (see Plate V-27).

Incidence and Prevalence About 13,200 cases of myeloma were diagnosed in 2000, and 11,200 people died from the disease. Myeloma increases in incidence with age. The median age at diagnosis is 68 years; it is rare under age 40. The yearly incidence is around 4 per 100,000 and remarkably similar throughout the world. Males are slightly more commonly affected than females, and blacks have nearly twice the incidence of whites. In the age group over 25 the incidence is 30 per 100,000. Myeloma accounts for about 1% of all malignancies in whites and 2% in blacks; 13% of all hematologic cancers in whites and 33% in blacks.

Pathogenesis and Clinical Manifestations (Table 113-1) Bone pain is the most common symptom in myeloma, affecting nearly 70% of patients. The pain usually involves the back and ribs, and unlike the pain of metastatic carcinoma, which often is worse at night, the pain of myeloma is precipitated by movement. Persistent localized pain in a patient with myeloma usually signifies a pathologic fracture. The bone lesions of myeloma are caused by the proliferation of tumor cells and the activation of osteoclasts that destroy the bone. The osteoclasts respond to osteoclast activating factors (OAF) made by the myeloma cells [OAF activity can be mediated by several cytokines, includingIL-1, lymphotoxin, and tumor necrosis factor (TNF)]. However, production of these factors stops following administration of glucocorticoids or interferon (IFN)-a. The bone lesions are lytic in nature and are rarely associated with osteoblastic new bone formation. Therefore, radioisotopic bone scanning is less useful in diagnosis than is plain radiography. The bony lysis results in substantial mobilization of calcium

from bone, and serious acute and chronic complications of hypercalcemia may dominate the clinical picture (see below). Localized bone lesions may expand to the point that mass lesions may be palpated, especially on the skull (Fig. 113-3), clavicles, and sternum, and the collapse of vertebrae may lead to spinal cord compression.

The next most common clinical problem in patients with myeloma is susceptibility to bacterial infections. The most common infections are pneumonias and pyelonephritis, and the most frequent pathogens are *Streptococcus pneumoniae*, *Staphylococcus aureus*, and *Klebsiella pneumoniae* in the lungs and *Escherichia coli* and other gram-negative organisms in the urinary tract. In about 25% of patients, recurrent infections are the presenting features, and over 75% of patients will have a serious infection at some time in their course. The susceptibility to infection has several contributing causes. First, patients with myeloma have diffuse hypogammaglobulinemia if the M component is excluded. The hypogammaglobulinemia is related to both decreased production and increased destruction of normal antibodies. Moreover, some patients generate a population of circulating regulatory cells in response to their myeloma that can suppress normal antibody synthesis. In the case of IgG myeloma, normal IgG antibodies are broken down more rapidly than normal because the catabolic rate for IgG antibodies varies directly with the serum concentration. The large M component results in fractional catabolic rates of 8 to 16% instead of the normal 2%. These patients have very poor antibody responses, especially to polysaccharide antigens such as those on bacterial cell walls. Most measures of T cell function in myeloma are normal, but a subset of CD4+ cells may be decreased. Granulocyte lysozyme content is low, and granulocyte migration is not as rapid as normal in patients with myeloma, probably the result of a tumor product. There are also a variety of abnormalities in complement functions in myeloma patients. All these factors contribute to the immune deficiency of these patients.

Renal failure occurs in nearly 25% of myeloma patients, and some renal pathology is noted in over half. Many factors contribute to this. Hypercalcemia is the most common cause of renal failure. Glomerular deposits of amyloid, hyperuricemia, recurrent infections, and occasional infiltration of the kidney by myeloma cells all may contribute to renal dysfunction. However, tubular damage associated with the excretion of light chains is almost always present. Normally, light chains are filtered, reabsorbed in the tubules, and catabolized. With the increase in the amount of light chains presented to the tubule, the tubular cells become overloaded with these proteins, and tubular damage results either directly from light chain toxic effects or indirectly from the release of intracellular lysosomal enzymes. The earliest manifestation of this tubular damage is the adult Fanconi syndrome (a type 2 proximal renal tubular acidosis), with increased loss of glucose, amino acids, and defects in the ability of the kidney to acidify and concentrate the urine. The proteinuria is not accompanied by hypertension, and the protein is nearly all light chains. Generally, very little albumin is in the urine because glomerular function is usually normal. When the glomeruli are involved, the proteinuria is nonselective. Patients with myeloma also have a decreased anion gap [i.e., $Na^+ - (Cl^- + HCO_3^-)$] because the M component is cationic, resulting in retention of chloride. This is often accompanied by hyponatremia that is felt to be artificial (pseudohyponatremia) because each volume of serum has less water as a result of the increased protein. Myeloma patients are susceptible to developing acute renal failure if they become dehydrated.

Anemia occurs in about 80% of myeloma patients. It is usually normocytic and normochromic and related both to the replacement of normal marrow by expanding tumor cells and to the inhibition of hematopoiesis by factors made by the tumor. In addition, mild hemolysis may contribute to the anemia. A larger than expected fraction of patients may have megaloblastic anemia due to either folate or vitamin B_{12}deficiency. Granulocytopenia and thrombocytopenia are very rare. Clotting abnormalities may be seen due to the failure of antibody-coated platelets to function properly or to the interaction of the M component with clotting factors I, II, V, VII, or VIII. Raynaud's phenomenon and impaired circulation may result if the M component forms cryoglobulins, and hyperviscosity syndromes may develop depending on the physical properties of the M component (most common with IgM, IgG3, and IgA paraproteins). Hyperviscosity is defined on the basis of the relative viscosity of serum as compared with water. Normal relative serum viscosity is 1.8 (i.e., serum is normally almost twice as viscous as water). Symptoms of hyperviscosity occur at a level of 5 to 6, a level usually reached at paraprotein concentrations of around 40 g/L (4 g/dL) for IgM, 50 g/L (5 g/dL) for IgG3, and 70 g/L (7 g/dL) for IgA.

Although neurologic symptoms occur in a minority of patients, they may have many causes. Hypercalcemia may produce lethargy, weakness, depression, and confusion. Hyperviscosity may lead to headache, fatigue, visual disturbances, and retinopathy. Bony damage and collapse may lead to cord compression, radicular pain, and loss of bowel and bladder control. Infiltration of peripheral nerves by amyloid can be a cause of carpal tunnel syndrome and other sensorimotor mono- and polyneuropathies.

Many of the clinical features of myeloma, e.g., cord compression, pathologic fractures, hyperviscosity, sepsis, and hypercalcemia, can present as medical emergencies. Despite the widespread distribution of plasma cells in the body, tumor expansion is dominantly within bone and bone marrow and, for reasons unknown, rarely causes enlargement of spleen, lymph nodes, or gut-associated lymphatic tissue.

Diagnosis and Staging The classic triad of myeloma is marrow plasmacytosis (>10%), lytic bone lesions, and a serum and/or urine M component. The diagnosis may be made in the absence of bone lesions if the plasmacytosis is associated with a progressive increase in the M component over time or if extramedullary mass lesions develop. There are two important variants of myeloma, solitary bone plasmacytoma and extramedullary plasmacytoma. These lesions are associated with an M component in fewer than 30% of the cases, they may affect younger individuals, and both are associated with median survivals of 10 or more years. Solitary bone plasmacytoma is a single lytic bone lesion without marrow plasmacytosis. Extramedullary plasmacytomas usually involve the submucosal lymphoid tissue of the nasopharynx or paranasal sinuses without marrow plasmacytosis. Both tumors are highly responsive to local radiation therapy. If an M component is present, it should disappear after treatment. Solitary bone plasmacytomas may recur in other bony sites or evolve into myeloma. Extramedullary plasmacytomas rarely recur or progress.

The most difficult differential diagnosis in patients with myeloma involves their separation from individuals with benign monoclonal gammopathies or monoclonal gammopathies of uncertain significance (MGUS). MGUS are vastly more common than

myeloma, occurring in 1% of the population over age 50 and in up to 10% over age 75. Patients with MGUS usually have <10% bone marrow plasma cells;<30 g/L (3 g/dL) of M components; no urinary Bence Jones protein; and no anemia, renal failure, lytic bone lesions, or hypercalcemia. When bone marrow cells are exposed to radioactive thymidine in order to quantitate dividing cells, patients with MGUS always have a labeling index <1%; patients with myeloma always have a labeling index >1%. Other discriminators include plasma cell acid phosphatase and b-glucuronidase, both of which are low in MGUS patients, and the salmon calcitonin stimulation test, which is positive only in patients with active ongoing bone destruction. With long-term follow-up, about 25% of patients with MGUS go on to develop myeloma. Typically, patients with MGUS require no therapy. Their survival is about 2 years shorter than age-matched controls without MGUS.

The clinical evaluation of patients with myeloma includes a careful physical examination searching for tender bones and masses. It is paradoxical that only a small minority of patients have an enlargement of the spleen and lymph nodes, the physiologic sites of antibody production. Chest and bone radiographs may reveal lytic lesions or diffuse osteopenia. A complete blood count with differential may reveal anemia. Erythrocyte sedimentation rate is elevated. Rare patients (~2%) may have plasma cell leukemia with more than 2000 plasma cells/uL. This may be seen in disproportionate frequency in IgD (12%) and IgE (25%) myelomas. Serum calcium, urea nitrogen, creatinine, and uric acid levels may be elevated. Protein electrophoresis and measurement of serum immunoglobulins are useful for detecting and characterizing M spikes, supplemented by immunoelectrophoresis, which is especially sensitive for identifying low concentrations of M components not detectable by protein electrophoresis. A 24-h urine specimen is necessary to quantitate protein excretion, and a concentrated aliquot is used for electrophoresis and immunologic typing of any M component. Serum alkaline phosphatase is usually normal even with extensive bone involvement because of the absence of osteoblastic activity. It is also important to quantitate serum b_2-microglobulin (see below). Serum soluble IL-6 receptor levels and C-reactive protein may reflect physiologic IL-6 levels in the patient.

The serum M component will be IgG in 53% of patients, IgA in 25%, and IgD in 1%; 20% of patients will have only light chains in serum and urine. Dipsticks for detecting proteinuria are not reliable at identifying light chains, and the heat test for detecting Bence Jones protein is falsely negative in about 50% of patients with light chain myeloma. Fewer than 1% of patients have no identifiable M component; these patients usually have light chain myelomas in which renal catabolism has made the light chains undetectable in the urine. IgD myeloma may also present as light chain myeloma. About two-thirds of patients with serum M components also have urinary light chains. The light chain isotype may have an impact on survival. Patients secreting lambda light chains have a significantly shorter overall survival than those secreting kappa light chains. It is not clear whether this is due to some genetically important determinant of cell proliferation or because lambda light chains are more likely to cause renal damage and form amyloid than are kappa light chains. The heavy chain isotype may have an impact on patient management as well. About half of patients with IgM paraproteins develop hyperviscosity compared with only 2 to 4% of patients with IgA and IgG M components. Among IgG myelomas, it is the IgG3 subclass that has the highest tendency to form both concentration- and temperature-dependent aggregates, leading to hyperviscosity

and cold agglutination at lower serum concentrations.

The staging system for patients with myeloma is a functional system for predicting survival and is based on a variety of clinical and laboratory tests, unlike the anatomic staging systems for solid tumors. Details of the staging system are given in Table 113-2. Based on the hemoglobin, calcium, M component, and degree of skeletal involvement, the total-body tumor burden is estimated to be low (stage I, $<0.6 \times 10^{12}$ cells/m^2), intermediate (stage II, 0.6 to 1.2×10^{12} cells/m^2), or high (stage III, $>1.2 \times 10^{12}$ cells/m^2), and the stages are further subdivided on the basis of renal function [A if serum creatinine <177 mol/L (<2 mg/dL), B if >177 (>2)]. Patients in stage IA have a median survival of more than 5 years and those in stage IIIB about 15 months. β_2-Microglobulin is a protein of 11,000 mol wt with homologies with the constant region of immunoglobulins that is the light chain of the class I major histocompatibility antigens (HLA-A, -B, -C) on the surface of every cell. Serum β_2-microglobulin is the single most powerful predictor of survival and can substitute for staging. Patients with β_2-microglobulin levels <0.004 g/L have a median survival of 43 months and those with levels >0.004 g/L only 12 months. It is also felt that once the diagnosis of myeloma is firm, histologic features of atypia may also exert an influence on prognosis. IL-6 may be an autocrine and/or paracrine growth factor for myeloma cells; elevated levels are associated with more aggressive disease. High labeling index and high levels of lactate dehydrogenase and thymidine kinase are also associated with poor prognosis.

Other factors that may influence prognosis are the number of cytogenetic abnormalities, % plasma cells in the marrow, performance status, and serum levels of IL-6, soluble IL-6 receptors, C-reactive protein, hepatocyte growth factor, C-terminal cross-linked telopeptide of collagen I, TGF-b, and syndecan-1.

TREATMENT

About 10% of patients with myeloma will have an indolent course demonstrating only very slow progression of disease over many years. Such patients only require antitumor therapy when the serum myeloma protein level rises above 50 g/L (5 g/dL) or progressive bone lesions develop. Patients with solitary bone plasmacytomas and extramedullary plasmacytomas may be expected to enjoy prolonged disease-free survival after local radiation therapy to a dose of around 40 Gy. There is a low incidence of occult marrow involvement in patients with solitary bone plasmacytoma. Such patients are usually detected because their serum M component falls slowly or disappears initially only to return after a few months. These patients respond well to systemic chemotherapy.

The vast majority of patients with myeloma require therapeutic intervention. In general such therapy is of two sorts: systemic chemotherapy to control the progression of myeloma, and symptomatic supportive care to prevent serious morbidity from the complications of the disease. All patients with stage II or III disease and stage I patients exhibiting Bence Jones proteinuria, progressive lytic bone lesions, vertebral compression fractures, recurrent infections, or rising serum M component should be treated with systemic combination chemotherapy. Therapy can prolong and improve the quality of life for myeloma patients.

The standard treatment has consisted of intermittent pulses of an alkylating agent [L-phenylalanine mustard (L-PAM, melphalan), cyclophosphamide, or chlorambucil] and prednisone administered for 4 to 7 days every 4 to 6 weeks. The alkylating agents appear to be roughly equally active, but resistance to one agent is often accompanied by resistance to the others. The usual doses are as follows: melphalan, 8 mg/m^2 per day; cyclophosphamide, 200 mg/m^2 per day; chlorambucil, 8 mg/m^2 per day; prednisone, 25 to 60 mg/m^2 per day. Melphalan is used most commonly, but because of their near equivalence in antitumor efficacy, we favor cyclophosphamide as the alkylating agent because it is less toxic to the marrow stem cell compartment and results in a lower incidence of acute myelodysplastic syndromes than do the other alkylating agents. Doses may need adjustment based on marrow tolerance. However, there are few constraints on the dose of the steroid pulse, and it appears that more is better. Patients responding to therapy generally have a prompt and gratifying reduction in bone pain, hypercalcemia, and anemia, and often have fewer infections. The serum M component lags substantially behind the symptomatic improvement, often taking 4 to 6 weeks to fall. This fall depends on the rate of tumor kill and the fractional catabolic rate of immunoglobulin, which in turn depends on the serum concentration (for IgG). Light chain excretion, with a functional half-life of approximately 6 h, may fall within the first week of treatment. However, since urine light chain levels may relate to renal tubular function, they are not a reliable measure of tumor cell kill. Calculations of tumor cell kill are made by extrapolation of the serum M component level and rely heavily on the assumption that every tumor cell produces immunoglobulin at a constant rate. About 60% of patients will achieve at least a 75% reduction in serum M component level and tumor cell mass in response to an alkylating agent and prednisone. Although this is a tumor reduction of less than one log, clinical responses may last many months. The important feature of the level of the M protein is not how far or how fast it falls, but the rate of its increase after therapy. Efforts to improve the fraction of patients responding and the degree of response have involved adding other active chemotherapeutic agents to the treatment program. Patients with more advanced disease may benefit most from such an approach. High-dose therapy with hematopoietic support is also being tested in younger patients. Sequential treatment with combination chemotherapy regimens followed by two successive high-dose melphalan treatments, each supported with peripheral blood stem cell transplants, have achieved complete responses in 50% of patients treated within a year of diagnosis. Complete responses are rare (<10%) with standard therapy. Long-term follow-up is not yet available. Allogeneic transplants may also produce high response rates, but treatment-related mortality may be as high as 40%.

The ideal duration of therapy has not been determined. Most physicians treat every 4 to 6 weeks for 1 or 2 years. Cessation of therapy is followed by relapse, usually within a year. Retreatment may be associated with a second response in up to 80% of patients. Maintenance therapy (e.g., with IFN-a) may prolong the duration of response, but this therapy is toxic and has generally not prolonged survival. The regrowth rate of the tumor during relapse accelerates with each relapse. This observation suggests that kinetic resistance to therapy (i.e., increase in cycling cells) is perhaps more important than drug resistance controlled by mdr-1 expression. Patients often respond to treatment, but the length of the response progressively shortens. Patients primarily resistant to initial therapy have a median survival of less than a year. High-dose pulsed steroids used alone (200 mg prednisone every other day or 1 g/m^2 per day methylprednisolone for 5

days) or VAD combination chemotherapy (vincristine, 0.4 mg/d in a 4-day continuous infusion; doxorubicin, 9 mg/m^2 per day in a 4-day continuous infusion; dexamethasone, 40 mg/d for 4 days per week for 3 weeks) may offer useful palliation in patients resistant to primary therapy. High-dose melphalan has activity in patients with refractory disease. Thalidomide, which inhibits angiogenesis, also produces responses in refractory cases, but at doses that may cause somnolence.

About 15% of patients die within the first 3 months after diagnosis; subsequently, the death rate is about 15% per year. The disease usually follows a chronic course for 2 to 5 years before developing an acute terminal phase, usually marked by the development of pancytopenia with a cellular marrow that is refractory to treatment. Widespread organ infiltration by myeloma cells occurs, and survival is less than 6 months. About 46% of patients die in the chronic phase of disease from progressive myeloma (16%) and renal failure (10%), sepsis (14%), or both (6%). Death in the acute terminal phase (26%) is chiefly from progressive myeloma (13%) and sepsis (9%). Five percent of patients die of acute leukemia, myeloblastic or monocytic. Although it has been debated that this is related to the primary disease, it appears more likely to be the result of chronic therapy with alkylating agents. Nearly 23% of patients die of myocardial infarction, chronic lung disease, diabetes, or stroke, all intercurrent illnesses related more to the age of the patient group than to the tumor.

Supportive care directed at the anticipated complications of the disease may be as important as primary antitumor therapy. The hypercalcemia generally responds well to glucocorticoid therapy, hydration, and natriuresis. Calcitonin may add to the inhibitory effects of steroids on bone resorption. Bisphosphonates (e.g., pamidronate 90 mg once a month) reduce osteoclastic bone resorption and preserve performance status and quality of life; antitumor effects are also possible. Treatments aimed at strengthening the skeleton, such as fluorides, calcium, and vitamin D, with or without androgens, have been suggested but are not of proven efficacy. Iatrogenic worsening of renal function may be prevented by the use of allopurinol during chemotherapy to avoid urate nephropathy and by maintaining a high fluid intake to prevent dehydration and to help excrete light chains and calcium. In the event of acute renal failure, plasmapheresis is approximately 10 times more effective at clearing light chains than peritoneal dialysis, and acutely reducing the protein load may result in functional improvement. Urinary tract infections should be watched for and treated early. Chronic dialysis probably should not be initiated in patients who have failed to respond to antitumor therapy. Plasmapheresis may be the treatment of choice for hyperviscosity syndromes. Although the pneumococcus is a dreaded pathogen in myeloma patients, pneumococcal polysaccharide vaccines may not elicit an antibody response. The advent of intravenous gamma globulin preparations raises some hope that prophylactic administration may prevent some serious infections, but this has not been tested. Chronic oral antibiotic prophylaxis is probably not warranted. Patients developing neurologic symptoms in the lower extremities, severe localized back pain, or problems with bowel and bladder control may need emergency myelography and radiation therapy for palliation. Most bone lesions respond to analgesics and chemotherapy, but certain painful lesions may respond most promptly to localized radiation. The chronic anemia may respond to hematinics (iron, folate, cobalamin), and some have responded to androgens. The pathogenesis of the anemia should be established and specific therapy instituted, where possible.

WALDENSTROM'S MACROGLOBULINEMIA

In 1948, Waldenstrom described a malignancy of lymphoplasmacytoid cells that secreted IgM. In contrast to myeloma, the disease was associated with lymphadenopathy and hepatosplenomegaly, but the major clinical manifestation was the hyperviscosity syndrome. The disease resembles the related diseases chronic lymphocytic leukemia, myeloma, and lymphocytic lymphoma. Waldenstrom's macroglobulinemia and IgM myeloma both follow a similar clinical course. The diagnosis of IgM myeloma is usually reserved for patients with lytic bone lesions and is important only because of the hazard of pathologic fractures.

The cause of macroglobulinemia is unknown. The disease is similar to myeloma in being slightly more common in men and occurring with increased incidence with age (median 64 years). There have been reports that the IgM in some patients with macroglobulinemia may have specificity for myelin-associated glycoprotein (MAG), a protein that has been associated with demyelinating disease of the peripheral nervous system and may be lost earlier and to a greater extent than the better known myelin basic protein in patients with multiple sclerosis. Sometimes patients with macroglobulinemia develop a peripheral neuropathy before the appearance of the neoplasm. There is speculation that the whole process begins with a viral infection that may elicit an antibody response that cross-reacts with a normal tissue component.

Like myeloma, the disease involves the bone marrow, but unlike myeloma, it does not cause bone lesions or hypercalcemia. Like myeloma, a serum M component is present in the serum in excess of 30 g/L (3 g/dL), but unlike myeloma, the size of the IgM paraprotein results in little renal excretion and only around 20% of patients excrete light chains. Therefore, renal disease is not common. The light chain isotype is kappa in 80% of the cases. Patients present with weakness, fatigue, and recurrent infections, similar to myeloma patients, but epistaxis, visual disturbances, and neurologic symptoms such as peripheral neuropathy, dizziness, headache, and transient paresis are much more common in macroglobulinemia. Physical examination reveals adenopathy and hepatosplenomegaly, and ophthalmoscopic examination may reveal vascular segmentation and dilatation of the retinal veins characteristic of hyperviscosity states. Patients may have a normocytic, normochromic anemia, but rouleaux formation and a positive Coombs' test are much more common than in myeloma. Malignant lymphocytes are usually present in the peripheral blood. About 10% of macroglobulins are cryoglobulins. These are pure M components and are not the mixed cryoglobulins seen in rheumatoid arthritis and other autoimmune diseases. Mixed cryoglobulins are composed of IgM or IgA complexed with IgG, for which they are specific. In both cases, Raynaud's phenomenon and serious vascular symptoms precipitated by the cold may occur, but mixed cryoglobulins are not commonly associated with malignancy. Patients suspected of having a cryoglobulin based on history and physical examination should have their blood drawn into a warm syringe and delivered to the laboratory in a container of warm water to avoid errors in quantitating the cryoglobulin.

TREATMENT

Control of serious hyperviscosity symptoms such as an altered state of consciousness

or paresis can be achieved acutely by plasmapheresis because 80% of the IgM paraprotein is intravascular. Fludarabine (25 mg/$m_2$per day for 5 days every 4 weeks) or cladribine (0.1 mg/kg per day for 7 days every 4 weeks) are highly effective single agents. About 80% of patients respond to chemotherapy, and their median survival is over 3 years. The absence of other serious organ toxicities results in a longer life span of patients with macroglobulinemia compared with those with myeloma.

POEMS SYNDROME

The features of this syndrome are *p*olyneuropathy, *o*rganomegaly, *e*ndocrinopathy, *m*ultiple myeloma, and *s*kin changes (POEMS). Patients usually have a severe, progressive sensorimotor polyneuropathy associated with sclerotic bone lesions from myeloma. Polyneuropathy occurs in about 1.4% of myelomas, but the POEMS syndrome is only a rare subset of that group. Unlike typical myeloma, hepatomegaly and lymphadenopathy occur in about two-thirds of patients, and splenomegaly is seen in one-third. The lymphadenopathy frequently resembles Castleman's disease histologically, a condition that has been linked toIL-6 overproduction. The endocrine manifestations include amenorrhea in women and impotence and gynecomastia in men. Hyperprolactinemia due to loss of normal inhibitory control by the hypothalamus may be associated with other central nervous system manifestations such as papilledema and elevated cerebrospinal fluid pressure and protein. Type 2 diabetes mellitus occurs in about one-third of patients. Hypothyroidism and adrenal insufficiency are occasionally noted. Skin changes are diverse: hyperpigmentation, hypertrichosis, skin thickening, and digital clubbing. Other manifestations include peripheral edema, ascites, pleural effusions, fever, and thrombocytosis.

The pathogenesis of the disease is unclear, but high circulating levels of the proinflammatory cytokinesIL-1, IL-6, andTNF have been documented and levels of the inhibitory cytokine transforming growth factorb (TGF-b) are lower than expected. Treatment of the myeloma may result in an improvement in the other disease manifestations.

HEAVY CHAIN DISEASES

The heavy chain diseases are rare lymphoplasmacytic malignancies. Their clinical manifestations vary with the heavy chain isotype. Patients secrete a defective heavy chain that usually has an intact Fc fragment and a deletion in the Fd region. Gamma, alpha, and mu heavy chain diseases have been described, but no reports of delta or epsilon heavy chain diseases have appeared. Molecular biologic analysis of these tumors has revealed structural genetic defects that may account for the aberrant chain secreted.

Gamma Heavy Chain Disease (Franklin's Disease) This disease affects people of widely different age groups and countries of origin. It is characterized by lymphadenopathy, fever, anemia, malaise, hepatosplenomegaly, and weakness. Its most distinctive symptom is palatal edema, resulting from node involvement of Waldeyer's ring, and this may progress to produce respiratory compromise. The diagnosis depends on the demonstration of an anomalous serum M component [often<20 g/L (<2 g/dL)] that reacts with anti-IgG but not anti-light chain reagents. *The M*

component is typically present in both serum and urine. Most of the paraproteins have been of the gamma$_1$subclass, but other subclasses have been seen. The patients may have thrombocytopenia, eosinophilia, and nondiagnostic bone marrow. Patients usually have a rapid downhill course and die of infection; however, some patients have survived 5 years with chemotherapy.

Alpha Heavy Chain Disease (Seligmann's Disease) This is the most common of the heavy chain diseases. It is closely related to a malignancy known as *Mediterranean lymphoma*, a disease that affects young people in parts of the world where intestinal parasites are common, such as the Mediterranean, Asia, and South America. The disease is characterized by an infiltration of the lamina propria of the small intestine with lymphoplasmacytoid cells that secrete truncated alpha chains. Demonstrating alpha heavy chains is difficult because the alpha chains tend to polymerize and appear as a smear instead of a sharp peak on electrophoretic profiles. Despite the polymerization, hyperviscosity is not a common problem in alpha heavy chain disease. Without J chain-facilitated dimerization, viscosity does not increase dramatically. Light chains are absent from serum and urine. The patients present with chronic diarrhea, weight loss, and malabsorption and have extensive mesenteric and para-aortic adenopathy. Respiratory tract involvement occurs rarely. Patients may vary widely in their clinical course. Some may develop diffuse aggressive histologies of malignant lymphoma. Chemotherapy may produce long-term remissions. Rare patients appear to have responded to antibiotic therapy, raising the question of the etiologic role of antigenic stimulation, perhaps by some chronic intestinal infection. Chemotherapy plus antibiotics may be more effective than chemotherapy alone.

Mu Heavy Chain Disease The secretion of isolated mu heavy chains into the serum appears to occur in a very rare subset of patients with chronic lymphocytic leukemia. The only features that may distinguish patients with mu heavy chain disease are the presence of vacuoles in the malignant lymphocytes and the excretion of kappa light chains in the urine. The diagnosis requires ultracentrifugation or gel filtration to confirm the nonreactivity of the paraprotein with the light chain reagents, because some intact macroglobulins fail to interact with these serums. The tumor cells seem to have a defect in the assembly of light and heavy chains, because they appear to contain both in their cytoplasm. There is no evidence that such patients should be treated differently from other patients with chronic lymphocytic leukemia (Chap. 112).

(Bibliography omitted in Palm version)

Back to Table of Contents

114. TRANSFUSION BIOLOGY AND THERAPY - *Jeffery S. Dzieczkowski*, *Kenneth C. Anderson*

BLOOD GROUP ANTIGENS AND ANTIBODIES

The study of red blood cell (RBC) antigens and antibodies forms the foundation of transfusion medicine. Serologic studies initially characterized these antigens, but now the molecular composition and structure of many are known. Antigens, either carbohydrate or protein, are assigned to a blood group system based upon the structure and the similarity of the determinant epitopes. Other cellular blood elements and plasma proteins are also antigenic and can result in *alloimmunization*, the production of antibodies directed against the blood group antigens of another individual. These antibodies are called *alloantibodies*.

Antibodies directed againstRBCantigens may result from "natural" exposure, particularly to carbohydrates that mimic some blood group antigens. Those antibodies that occur via natural stimuli are usually produced by a T cell-independent response (thus, generating no memory) and are IgM isotype. *Autoantibodies* (antibodies against autologous blood group antigens) arise spontaneously or as the result of infectious sequelae (e.g., from *Mycoplasma pneumoniae*) and are also often IgM. These antibodies are often clinically insignificant due to their low affinity for antigen at body temperature. However, IgM antibodies can activate the complement cascade and result in hemolysis. Antibodies that result from allogeneic exposure, such as transfusion or pregnancy, are usually IgG. IgG antibodies commonly bind to antigen at warmer temperatures and may hemolyze RBCs. Unlike IgM antibodies, IgG antibodies can cross the placenta and bind fetal erythrocytes bearing the corresponding antigen, resulting in hemolytic disease of the newborn, or *hydrops fetalis*.

Alloimmunization to leukocytes, platelets, and plasma proteins may also result in transfusion complications such as fevers and urticaria but generally does not cause hemolysis. Assay for these other alloantibodies is not routinely performed; however, they may be detected using special assays.

ABO ANTIGENS AND ANTIBODIES

The first blood group antigen system, recognized in 1900, was ABO, the most important in transfusion medicine. The major blood groups of this system are A, B, AB, and O. O typeRBCs lack A or B antigens. These antigens are carbohydrates attached to a precursor backbone, may be found on the cellular membrane either as glycosphingolipids or glycoproteins, and are secreted into plasma and body fluids as glycoproteins. H substance is the immediate precursor upon which the A and B antigens are added. This H substance is formed by the addition of fucose to the glycolipid or glycoprotein backbone. The subsequent addition of *N*-acetylgalactosamine creates the A antigen, while the addition of galactose produces the B antigen.

The genes that determine the A and B phenotypes are found on chromosome 9p and are expressed in a Mendelian codominant manner. The gene products are glycosyl transferases, which confer the enzymatic capability of attaching the specific antigenic carbohydrate. Individuals who lack the "A" and "B" transferases are phenotypically type

"O," while those who inherit both transferases are type "AB." Rare individuals lack the H gene, which codes for fucose transferase, and cannot form H substance. These individuals are homozygous for the silent h allele (hh) and have Bombay phenotype (O_h).

The ABO blood group system is important because essentially all individuals produce antibodies to the ABH carbohydrate antigen that they lack. The naturally occurring anti-A and anti-B antibodies are termed *isoagglutinins*. Thus, type A individuals produce anti-B, while type B individuals make anti-A. Neither isoagglutinin is found in type AB individuals, while type O individuals produce both anti-A and anti-B. Thus, persons with type AB are "universal recipients" because they do not have antibodies against any ABO phenotype, while persons with type O blood can donate to essentially all recipients because their cells are not recognized by any ABO isoagglutinins. The rare individuals with Bombay phenotype produce antibodies to H substance (which is present on all red cells except those of hh phenotype) as well as to both A and B antigens and are therefore compatible only with other hh donors.

In most people, A and B antigens are secreted by the cells and are present in the circulation. Nonsecretors are susceptible to a variety of infections (e.g., *Candida albicans*, *Neisseria meningitidis*, *Streptococcus pneumoniae*, *Haemophilus influenzae*) as many organisms may bind to polysaccharides on cells. Soluble blood group antigens may block this binding.

RH SYSTEM

The Rh system is the second most important blood group system in pretransfusion testing. The Rh antigens are found on a 30- to 32-kDaRBCmembrane protein, which has no defined function. Although more than 40 different antigens in the Rh system have been described, five determinants account for the vast majority of phenotypes. The presence of the D antigen confers Rh "positivity," while people who lack the D antigen are Rh negative. Two allelic antigen pairs, E/e and C/c, are also found on the Rh protein. The three Rh genes, E/e, D, and C/c, are arranged in tandem on chromosome 1 and inherited as a haplotype, i.e., cDE or Cde. Two haplotypes can result in the phenotypic expression of two to five Rh antigens.

The D antigen is a potent alloantigen. About 15% of people lack this antigen. Exposure of these Rh-negative people to even small amounts of Rh-positive cells, by either transfusion or pregnancy, can result in the production of anti-D alloantibody.

OTHER BLOOD GROUP SYSTEMS AND ALLOANTIBODIES

More than 100 blood group systems are recognized, composed of more than 500 antigens. The presence or absence of certain antigens has been associated with various diseases and anomalies; antigens also act as receptors for infectious agents. Alloantibodies of importance in routine clinical practice are listed inTable 114-1.

Antibodies to *Lewis system* carbohydrate antigens are the most common cause of incompatibility during pretransfusion screening. The Lewis gene product is a fucosyl transferase and maps to chromosome 19. The antigen is not an integral membrane

structure but is adsorbed to the RBC membrane from the plasma. Antibodies to Lewis antigens are usually IgM and cannot cross the placenta. Lewis antigens may be adsorbed onto tumor cells and may be targets of therapy.

I system antigens are also oligosaccharides related to H, A, B, and Le. I and i are not allelic pairs but are carbohydrate antigens that differ only in the extent of branching. The i antigen is an unbranched chain that is converted by the I gene product, a glycosyltransferase, into a branched chain. The branching process affects all the ABH antigens, which become progressively more branched in the first 2 years of life. Some patients with cold agglutinin disease or lymphomas can produce anti-I autoantibodies that cause RBC destruction. Occasional patients with mononucleosis or *Mycoplasma* pneumonia may develop cold agglutinins of either anti-I or anti-i specificity. Most adults lack i expression; thus, finding a donor for patients with anti-i is not difficult. Even though most adults express I antigen, binding is generally low at body temperature. Thus, administration of warm blood prevents isoagglutination.

The *P system* is another group of carbohydrate antigens controlled by specific glycosyltransferases. Its clinical significance is in rare cases of syphilis and viral infection that lead to paroxysmal cold hemoglobinuria. In these cases, an unusual autoantibody to P is produced that binds to RBCs in the cold and fixes complement upon warming. Antibodies with these biphasic properties are called *Donath-Landsteiner antibodies*. The P antigen is also expressed on urothelial cells and may be a receptor for *Escherichia coli* binding.

The *MNSsU system* is regulated by genes on chromosome 4. M and N are determinants on glycophorin A, an RBC membrane protein, and S and s are determinants on glycophorin B. Anti-S and anti-s IgG antibodies may develop after pregnancy or transfusion and lead to hemolysis. Anti-U antibodies are rare but problematic; virtually every donor is incompatible because nearly all persons express U.

The *Kell* protein is very large (720 amino acids) and its secondary structure contains many different antigenic epitopes. The immunogenicity of Kell is third behind the ABO and Rh systems. The absence of the Kell precursor protein (controlled by a gene on X) is associated with acanthocytosis, shortened RBC survival, and a progressive form of muscular dystrophy that includes cardiac defects. This rare condition is called the *McLeod phenotype*. The K_x gene is linked to the 91-kDa component of the NADPH-oxidase on the X chromosome, deletion or mutation of which accounts for about 60% of cases of chronic granulomatous disease.

The *Duffy* antigens are codominant alleles, Fy_a and Fy_b, that also serve as receptors for *Plasmodium vivax*. More than 70% of persons in malaria-endemic areas lack these antigens, probably from selective influences of the infection on the population.

The *Kidd* antigens, Jk_a and Jk_b, may elicit antibodies transiently. A delayed hemolytic transfusion reaction that occurs with blood tested as compatible is often related to delayed appearance of anti-Jk_a.

PRETRANSFUSION TESTING

Pretransfusion testing of a potential recipient consists of the "type and screen." The "forward type" determines the ABO and Rh phenotype of the recipient'sRBC by using antisera directed against the A, B, and D antigens. The "reverse type" detects isoagglutinins in the patient's serum and should correlate with the ABO phenotype, or forward type.

The alloantibody screen identifies antibodies directed against otherRBCantigens. The alloantibody screen is performed by mixing patient serum with type O RBCs that contain the major antigens of most blood group systems and whose extended phenotype is known. The specificity of the alloantibody is identified by correlating the presence or absence of antigen with the results of the agglutination.

Cross matching is ordered when there is a high probability that the patient will require a packed RBC (PRBC) transfusion. Blood selected for cross matching must be ABO compatible and lack antigens for which the patient has alloantibodies. Nonreactive cross matching confirms the absence of any major incompatibility and reserves that unit for the patient.

In the case of Rh-negative patients, every attempt must be made to provide Rh-negative blood components to prevent alloimmunization to the D antigen. In an emergency, Rh-positive blood can be safely transfused to a Rh-negative patient who lacks anti-D; however, the recipient is likely to become alloimmunized and produce anti-D. Rh-negative women of child-bearing age who are transfused with products containing Rh-positiveRBCs should receive passive immunization with anti-D (RhoGam or WinRho) to reduce or prevent sensitization.

BLOOD COMPONENTS

Blood products intended for transfusion are routinely collected as whole blood (450 mL) in various anticoagulants. Most donated blood is processed into components:PRBCs, platelets, and fresh frozen plasma (FFP) or cryoprecipitate (Table 114-2). Whole blood is first separated into PRBCs and platelet-rich plasma by slow centrifugation. The platelet-rich plasma is then centrifuged at high speed to yield one unit of random donor (RD) platelets and one unit of FFP. Cryoprecipitate is produced by thawing FFP to precipitate the plasma proteins, which are then separated by centrifugation.

Apheresis technology is used for the collection of multiple units of platelets from a single donor. These single-donor apheresis platelets (SDAP) contain the equivalent of at least six units ofRDplatelets and have fewer contaminating leukocytes than pooled RD platelets.

Plasma may also be collected by apheresis. Plasma derivatives such as albumin, intravenous immunoglobulin, antithrombin, and coagulation factor concentrates are prepared from pooled plasma from many donors and are treated to eliminate infectious agents.

WHOLE BLOOD

Whole blood provides both oxygen-carrying capacity and volume expansion. It is the

ideal component for patients who have sustained acute hemorrhage of 25% or greater total blood volume loss. Whole blood is stored at 4°C to maintain erythrocyte viability, but platelet dysfunction and degradation of some coagulation factors occurs. In addition, 2,3-BPG levels fall over time, leading to an increase in the oxygen affinity of the hemoglobin and a decreased capacity to deliver oxygen to the tissues, a problem with all red cell storage. Whole blood is not readily available since it is routinely processed into components.

PACKED RED BLOOD CELLS

This product increases oxygen-carrying capacity in the anemic patient. Adequate oxygenation can be maintained with a hemoglobin content of 70 g/L in the normovolemic patient without cardiac disease; however, comorbid factors often necessitate transfusion at a higher threshold. The decision to transfuse should be guided by the clinical situation and not by an arbitrary laboratory value. In the critical care setting, liberal use of transfusions to maintain near normal levels of hemoglobin may have unexpected negative effects on survival. In most patients requiring transfusion, levels of hemoglobin of 100 g/L are sufficient to keep oxygen supply from being critically low.

PRBCs may be modified to prevent certain adverse reactions. Contaminating leukocytes are responsible for inducing fevers and causing alloimmunization to HLA antigens. Leukocytes can be removed by several methods. Bedside filtration is the most popular method and removes 99.9% of donor leukocytes. Leukoreduction may be done in the blood bank before storage of cellular components; this practice results in less cytokine release from the cells. Plasma, which may cause allergic reactions, can be removed from cellular blood components by washing.

PLATELETS

Thrombocytopenia is a risk factor for hemorrhage, and platelet transfusion reduces the incidence of bleeding. The threshold for prophylactic platelet transfusion is 10,000/uL. In patients without fever or infections, a threshold of 5000/uL may be sufficient to prevent spontaneous hemorrhage. For invasive procedures, 50,000/uL platelets is the usual target level.

Platelets are given either as pools prepared from five to eightRDs or as SDAPs from a single donor. In an unsensitized patient without increased platelet consumption [splenomegaly, fever, disseminated intravascular coagulation (DIC)], six to eight units of RD platelets (about 1 unit per 10 kg body weight) are transfused, and each unit is anticipated to increase the platelet count 5000 to 10,000/uL. Patients who have received multiple transfusions may be alloimmunized to many HLA- and platelet-specific antigens and have little or no increase in their posttransfusion platelet counts. Patients who may require multiple transfusions are best served by receiving SDAP and leukocyte-reduced components to lower the risk of alloimmunization.

Refractoriness to platelet transfusion may be evaluated using the corrected count increment (CCI):

where BSA is body surface area measured in square meters. The platelet count performed 1 h after the transfusion is acceptable if the CCI is 10 ´10^9/mL, and after 18 to 24 h an increment of 7.5 ´10^9/mL is expected. Patients who have suboptimal responses are likely to have received multiple transfusions and have antibodies directed against class I HLA antigens. Refractoriness can be investigated by detecting anti-HLA antibodies in the recipient's serum. Patients who are sensitized will often react with 100% of the lymphocytes used for the HLA-antibody screen, and HLA-matchedSDAPs should be considered for those patients who require transfusion. Although ABO-identical HLA-matched SDAPs provide the best chance for increasing the platelet count, locating these products is difficult. Platelet cross matching is available in some centers. Additional clinical causes for a low platelet CCI include fever, bleeding, splenomegaly,DIC, or medications in the recipient.

FRESH FROZEN PLASMA

FFPcontains stable coagulation factors and plasma proteins: fibrinogen, antithrombin, albumin, as well as proteins C and S. Indications for FFP include correction of coagulopathies, including the rapid reversal of coumadin; supplying deficient plasma proteins; and treatment of thrombotic thrombocytopenic purpura. FFP should not be routinely used to expand blood volume. FFP is an acellular component and does not transmit intracellular infections, e.g., cytomegalovirus (CMV). Patients who are IgA-deficient and require plasma support should receive FFP from IgA-deficient donors to prevent anaphylaxis (see below).

CRYOPRECIPITATE

Cryoprecipitate is a source of fibrinogen, factor VIII, and von Willebrand factor (vWF). It is ideal for supplying fibrinogen to the volume-sensitive patient. When factor VIII concentrates are not available, cyroprecipitate may be used since each unit contains approximately 80 units of factor VIII. Cryoprecipitate may also be used as a source of vWF for patients with dysfunctional (type II) or absent (type III) von Willebrand disease.

PLASMA DERIVATIVES

Plasma from thousands of donors may be pooled to derive specific protein concentrates, including albumin, intravenous immunoglobulin, antithrombin, and coagulation factors. In addition, donors who have high-titer antibodies to specific agents or antigens provide hyperimmune globulins, such as anti-D (RhoGam, WinRho), and antisera to hepatitis B virus (HBV), varicella-zoster virus,CMV, and other infectious agents.

ADVERSE REACTIONS TO BLOOD TRANSFUSION

Adverse reactions to transfused blood components occur despite multiple tests, inspections, and checks. Fortunately, the most common reactions are not life-threatening, although serious reactions can present with mild symptoms and signs. Some reactions can be reduced or prevented by modified (filtered, washed, or

irradiated) blood components. When an adverse reaction is suspected, the transfusion should be stopped and reported to the blood bank for investigation.

Transfusion reactions may result from immune and nonimmune mechanisms. Immune-mediated reactions are often due to preformed donor or recipient antibody; however, cellular elements may also cause adverse effects. Nonimmune causes of reactions are due to the chemical and physical properties of the stored blood component and its additives.

Infectious complications of transfusion have become less frequent, although fear of these complications remains a primary concern. The incidence of transfusion-related infections has been reduced substantially due to improved donor screening and testing of collected blood. Infections, like any adverse transfusion reaction, must be brought to the attention of the blood bank for appropriate studies (Table 114-3).

IMMUNE-MEDIATED REACTIONS

Acute Hemolytic Transfusion Reactions Immune-mediated hemolysis occurs when the recipient has preformed antibodies that lyse donor erythrocytes. The ABO isoagglutinins are responsible for the majority of these reactions, although alloantibodies directed against other RBC antigens, i.e., Rh, Kell, and Duffy, may result in hemolysis.

Acute hemolytic reactions may present with hypotension, tachypnea, tachycardia, fever, chills, hemoglobinemia, hemoglobinuria, chest and/or flank pain, and discomfort at the infusion site. Monitoring the patient's vital signs before and during the transfusion is important to identify reactions promptly. When acute hemolysis is suspected, the transfusion must be stopped immediately, intravenous access maintained, and the reaction reported to the blood bank. A correctly labeled posttransfusion blood sample and any untransfused blood should be sent to the blood bank for analysis. The laboratory evaluation for hemolysis includes the measurement of serum haptoglobin, lactate dehydrogenase (LDH), and indirect bilirubin levels.

The immune complexes that result inRBClysis can cause renal dysfunction and failure. Diuresis should be induced with intravenous fluids and furosemide or mannitol. Tissue factor released from the lysed erythrocytes may initiateDIC. Coagulation studies including prothrombin time (PT), activated partial thromboplastin time (aPTT), fibrinogen, and platelet count should be monitored in patients with hemolytic reactions.

Errors at the patient's bedside, such as mislabeling the sample or transfusing the wrong patient, are responsible for the majority of these reactions. The blood bank investigation of these reactions includes examination of the pre- and posttransfusion samples for hemolysis and repeat typing of the patient samples; direct antiglobulin test (DAT), sometimes called the direct Coombs test, of the posttransfusion sample; repeating the cross matching of the blood component; and checking all clerical records for errors. DAT detects the presence of antibody or complement bound toRBCs in vivo.

Delayed Hemolytic and Serologic Transfusion Reactions Delayed hemolytic transfusion reactions (DHTRs) are not completely preventable. These reactions occur in patients previously sensitized toRBCalloantigens who have a negative alloantibody

screen due to low antibody levels. When the patient is transfused with antigen-positive blood, an anamnestic response results in the early production of alloantibody that binds donor RBCs. The alloantibody is detectable 1 to 2 weeks following the transfusion, and the posttransfusionDAT may become positive due to circulating donor RBCs coated with antibody or complement. The transfused, alloantibody-coated erythrocytes are cleared by the extravascular reticuloendothelial system. These reactions are detected most commonly in the blood bank when a subsequent patient sample reveals a positive alloantibody screen or a new alloantibody in a recently transfused recipient.

No specific therapy is usually required, although additionalRBCtransfusions may be necessary. Delayed serologic transfusion reactions (DSTR) are similar toDHTR, as the DATis positive and alloantibody is detected; however, RBC clearance is not increased.

Febrile Nonhemolytic Transfusion Reaction The most frequent reaction associated with the transfusion of cellular blood components is a febrile nonhemolytic transfusion reaction (FNHTR). These reactions are characterized by chills and rigors and a 1°C or greater rise in temperature. FNHTR is diagnosed when other causes of fever in the transfused patient are ruled out. Antibodies directed against donor leukocyte and HLA antigens may mediate these reactions; thus, multiply transfused patients and multiparous women are felt to be at increased risk. Although antibodies may be demonstrated in the recipient's serum, investigation is not routinely done because of the mild nature of most FNHTR. The use of leukocyte-reduced blood products may prevent or delay sensitization to leukocyte antigens and thereby reduce the incidence of these febrile episodes. Cytokines released from cells within stored blood components may mediate FNHTR; thus, leukoreduction before storage may prevent these reactions. The incidence and severity of these reactions can be decreased in patients with recurrent reactions by premedicating with acetaminophen or other antipyretic agents.

Allergic Reactions Urticarial reactions are related to plasma proteins found in transfused components. Mild reactions may be treated symptomatically by temporarily stopping the transfusion and administering antihistamines (diphenhydramine, 50 mg orally or intramuscularly). The transfusion may be completed after the signs and/or symptoms resolve. Patients with a history of allergic transfusion reaction should be premedicated with an antihistamine. Cellular components can be washed to remove residual plasma for the extremely sensitized patient.

Anaphylactic Reaction This severe reaction presents after transfusion of only a few milliliters of the blood component. Symptoms and signs include difficulty breathing, coughing, nausea and vomiting, hypotension, bronchospasm, loss of consciousness, respiratory arrest, and shock. Treatment includes stopping the transfusion, maintaining vascular access, and administering epinephrine (0.5 to 1.0 mL of 1:1000 dilution SQ). Glucocorticoids may be required in severe cases.

Patients who are IgA-deficient may be sensitized to this Ig class and are at risk for anaphylactic reactions associated with plasma transfusion. Individuals with severe IgA deficiency should therefore receive only IgA-deficient plasma and washed cellular blood components. Patients who have anaphylactic or repeated allergic reactions to blood components should be tested for IgA deficiency.

Graft-Versus-Host Disease Graft-versus-host disease (GVHD) is a frequent complication of allogeneic bone marrow transplantation, in which viable lymphocytes from donor marrow attack and cannot be eliminated by an immunodeficient host. Transfusion-related GVHD is mediated by donor T lymphocytes that recognize host HLA antigens as foreign and mount an immune response, which is manifested clinically by the development of fever, a characteristic cutaneous eruption, diarrhea, and liver function abnormalities. GVHD can also occur when blood components that contain viable T lymphocytes are transfused to immunodeficient recipients or to immunocompetent recipients who share HLA antigens with the donor (e.g., a family donor). In addition to the aforementioned clinical features of GVHD, transfusion-associated GVHD (TA-GVHD) is characterized by marrow aplasia and pancytopenia. TA-GVHD is highly resistant to treatment with immunosuppressive therapies, including glucocorticoids, cyclosporine, antithymocyte globulin, and ablative therapy followed by allogeneic bone marrow transplantation. Clinical manifestations appear at 8 to 10 days, and death occurs at 3 to 4 weeks posttransfusion.

TA-GVHDcan be prevented by irradiation of cellular components (minimum of 2500 cGy) before transfusion to patients at risk. Patients at risk for TA-GVHD include fetuses receiving intrauterine transfusions, selected immunocompetent (e.g., lymphoma patients) or immunocompromised recipients, recipients of donor units known to be from a blood relative, and recipients who have undergone marrow transplantation. Directed donations by family members should be discouraged (they are not less likely to transmit infection); lacking other options, the blood products from family members should always be irradiated.

Transfusion-Related Acute Lung Injury This uncommon reaction results from the transfusion of donor plasma that contains high titer anti-HLA antibodies that bind recipient leukocytes. The leukocytes aggregate in the pulmonary vasculature and release mediators that increase capillary permeability. The recipient develops symptoms of respiratory compromise and signs of noncardiogenic pulmonary edema, including bilateral interstitial infiltrates on chest x-ray. Treatment is supportive, and patients usually recover without sequelae. Testing the donor's plasma for anti-HLA antibodies can support this diagnosis. The implicated donors are frequently multiparous women, and transfusion of their plasma component should be avoided.

Posttransfusion Purpura This reaction presents as thrombocytopenia 7 to 10 days after platelet transfusion and occurs predominantly in women. Platelet-specific antibodies are found in the recipient's serum, and the most frequently recognized antigen is HPA-1a found on the platelet glycoprotein IIIa receptor. The delayed thrombocytopenia is due to the production of antibodies that react to both donor and recipient platelets. Additional platelet transfusions can worsen the thrombocytopenia and should be avoided. Treatment with intravenous immunoglobulin may neutralize the effector antibodies, or plasmapheresis can be used to remove the antibodies.

Alloimmunization A recipient may become alloimmunized to a number of antigens on cellular blood elements and plasma proteins. Alloantibodies toRBCantigens are detected during pretransfusion testing, and their presence may delay finding antigen-negative crossmatch-compatible products for transfusion. Women of child-bearing age who are sensitized to certain RBC antigens (i.e., D, c, E, Kell, or

Duffy) are at risk for bearing a fetus with hemolytic disease of the newborn. Matching for D antigen is the only pretransfusion selection test to prevent RBC alloimmunization.

Alloimmunization to antigens on leukocytes and platelets can result in refractoriness to platelet transfusions. Once alloimmunization has developed, HLA-compatible platelets from donors who share similar antigens with the recipient may be difficult to find. Hence, prudent transfusion practice is directed at preventing sensitization through the use of leukocyte-reduced cellular components, as well as limiting antigenic exposure by the judicious use of transfusions and use ofSDAPs.

NONIMMUNOLOGIC REACTIONS

Fluid Overload Blood components are excellent volume expanders, and transfusion may quickly lead to volume overload. Monitoring the rate and volume of the transfusion, along with the use of a diuretic, can minimize this problem.

Hypothermia Refrigerated (4°C) or frozen (-18°C or below) blood components can result in hypothermia when rapidly infused. Cardiac dysrhythmias can result from exposing the sinoatrial node to cold fluid. Use of an in-line warmer will prevent this complication.

Electrolyte Toxicity RBCleakage during storage increases the concentration of potassium in the unit. Neonates and patients in renal failure are at risk for hyperkalemia. Preventive measures, such as using fresh or washed RBCs, are warranted for neonatal transfusions because this complication can be fatal.

Citrate, commonly used to anticoagulate blood components, chelates calcium and thereby inhibits the coagulation cascade. Hypocalcemia, manifested by circumoral numbness and/or tingling sensation of the fingers and toes, may result from multiple rapid transfusions. Because citrate is quickly metabolized to bicarbonate, calcium infusion is seldom required in this setting. If calcium or any other intravenous infusion is necessary, it must be given through a separate intravenous line.

Iron Overload Each unit ofRBCs contains 200 to 250 mg of iron. Symptoms and signs of iron overload affecting endocrine, hepatic, and cardiac function are common after 100 units of RBCs have been transfused (total body iron load of 20 g). Preventing this complication by using alternative therapies (e.g., erythropoietin) and judicious transfusion is preferable and cost effective. Deferoxamine and other chelating agents are available, but the response is often suboptimal.

Hypotensive Reactions Transient hypotension may be noted among transfused patients who take angiotensin-converting enzyme (ACE) inhibitors. Since blood products contain bradykinin that is normally degraded by ACE, patients on ACE inhibitors may have increased bradykinin levels that cause hypotension. The blood pressure typically returns to normal without intervention.

Immunomodulation Transfusion of allogeneic blood is immunosuppressive. Multiply transfused renal transplant recipients are less likely to reject the graft. However, in postoperative settings and in cancer patients, immune suppression is dangerous. The

use of leukocyte-depleted cellular products may reduce the immunosuppression, though controlled data have not been obtained.

INFECTIOUS COMPLICATIONS

Viral Infections

Hepatitis C virus (HCV) The use of an improved screening test for HCV antibodies has reduced the incidence of posttransfusion HCV infection to 1 in 103,000 transfusions. Infection with HCV may be asymptomatic or lead to chronic active hepatitis, cirrhosis, and liver failure.

Hepatitis B Virus Transfusion-associatedHBVinfection has been reduced with improved donor selection and screening, along with increased vaccination of the donor and recipient population. However, some data suggest that HBV is more commonly transmitted by transfusion thanHCV. Vaccination of individuals who require long-term transfusion therapy can prevent this complication.

Hepatitis G virus (HGV) This hepatotropic virus is transmitted by transfusion. Infection with HGV results in no apparent adverse effects. Routine testing is not available and does not appear to be warranted.

Human Immunodeficiency Virus Type 1 Intensive donor screening and testing has dramatically reduced the risk of HIV-1 infection by blood transfusion. Donated blood is tested for HIV-1 p24 antigen. Two antigen-positive seronegative donors have been identified. The risk of HIV-1 infection per transfusion episode is 1 in 676,000. A specific assay to detect antibodies to HIV-2 is also performed on donated blood. No cases of HIV-2 infection have been reported in the United States since 1992, and only three donors have been found to have HIV-2 antibodies.

Cytomegalovirus This ubiquitous virus infects 50% or more of the general population and is transmitted by the infected "passenger" white blood cells found in transfusedPRBCs or platelet components. Donated blood is not routinely tested for serologic evidence of donor exposure, but assays can be performed to identifyCMV-seronegative donors, if needed. Alternatively, cellular components that are leukocyte-reduced have a decreased risk of transmitting CMV, regardless of the serologic status of the donor. Groups at risk for CMV infections include immunosuppressed patients, CMV-seronegative transplant recipients, and neonates; these patients should receive seronegative or leukocyte-depleted components.

Human T lymphotropic virus (HTLV) type I Assays to detect HTLV-I and -II are used to screen all donated blood. HTLV-1 is associated with adult T cell leukemia/lymphoma and tropical spastic paraparesis in a small percentage of infected persons (Chap. 191). The reported risk of HTLV-I infection via transfusion is 1 in 641,000 transfusion episodes. HTLV-II is not clearly associated with any disease.

Parvovirus B-19 Blood components and products derived from pooled plasma can transmit this virus, the etiologic agent of erythema infectiosum, or fifth disease, in children. Parvovirus B-19 shows tropism for erythroid precursors and inhibits both

erythrocyte production and maturation. Pure red cell aplasia, presenting either as acute aplastic crisis or chronic anemia with shortenedRBCsurvival, may occur in individuals with an underlying hematologic disease, such as sickle cell disease or thalassemia. The fetus of a seronegative woman is at risk for developing hydrops if infected with this virus.

Bacterial Contamination Most bacteria do not grow well at cold temperatures; thus,PRBCs and FFPare not common sources of bacterial contamination. However, some gram-negative bacteria, notably *Yersinia* and *Pseudomonas* species, can grow at 1° to 6°C. Platelet concentrates, which are stored at room temperature, are more likely to be contaminated with skin contaminants such as gram-positive organisms, including coagulase-negative staphylococci.

Recipients of transfusions contaminated with bacteria may develop fever and chills, which can progress to septic shock andDIC. These reactions may occur abruptly, within minutes of initiating the transfusion, or after several hours. The onset of symptoms and signs is often sudden and fulminant, which aids in differentiating bacterial contamination from a FNHTR. The reactions, particularly those related to gram-negative contaminants, are the result of infused endotoxins formed within the contaminated stored component.

When contaminated transfusions are suspected (i.e., when there is sudden development of shock), the transfusion must be stopped immediately. Therapy is directed at supporting the recipient's blood pressure, cardiac output, oxygenation, and renal function. The laboratory investigation should include cultures of any untransfused component, along with the routine blood bank clerical checks and serologic studies. Broad-spectrum antibiotic coverage should be started immediately and may be adjusted based on culture and sensitivity.

Parasites Various parasites including those causing malaria, babesiosis, and Chagas' disease can be transmitted by blood transfusion rarely. Geographical migration and travel of donors can shift the incidence of these rare infections. Because these infections can prove fatal, they should be considered in the transfused patient in the appropriate clinical setting.

ALTERNATIVES TO TRANSFUSION

Alternatives to allogeneic blood transfusions that avoid homologous donor exposures with attendant immunologic and infectious risks remain attractive. Autologous blood is the best option when transfusion is anticipated. However, the cost:benefit ratio of autologous transfusion remains high. No transfusion is a zero-risk event; clerical errors and bacterial contamination remain potential complications even with autologous transfusions. Additional methods of autologous transfusion in the surgical patient include preoperative hemodilution, recovery of shed blood from sterile surgical sites, and postoperative drainage collection. Directed or designated donation from friends and family of the potential recipient has not been safer than volunteer donor component transfusions. Such directed donations may in fact place the recipient at higher risk for complications such asGVHD and alloimmunization.

Oxygen-carrying blood substitutes, such as perfluorocarbons and aggregated

hemoglobin solution, are presently in various stages of clinical trials. Granulocyte- and granulocyte-macrophage colony stimulating factor (G- or GM-CSF) are clinically useful to hasten leukocyte recovery in patients with leukopenia related to high-dose chemotherapy. Erythropoietin stimulates erythrocyte production in patients with anemia of chronic renal failure and other conditions, thus avoiding or reducing the need for transfusion. This hormone can also stimulate erythropoiesis in the autologous donor to enable additional donation. Thrombopoietin, a cytokine that promotes megakaryocyte proliferation and maturation, is being tested for its ability to reduce the need for platelet transfusion.

(Bibliography omitted in Palm version)

Back to Table of Contents

115. BONE MARROW AND STEM CELL TRANSPLANTATION - *Frederick R. Appelbaum*

Bone marrow transplantation is the generic term used to describe the collection and transplantation of hematopoietic stem cells. The procedure is usually carried out for one of two purposes: (1) to replace an abnormal but nonmalignant lymphohematopoietic system with one from a normal donor, or (2) to treat malignancy by allowing the administration of higher doses of myelosuppressive therapy than would otherwise be possible. The use of bone marrow transplantation has been steadily increasing, both because of the its demonstrated effectiveness in selected diseases and because of increasing availability of donors. The International Bone Marrow Transplant Registry estimates that about 50,000 transplants were performed during 1999.

THE HEMATOPOIETIC STEM CELL

Several features of the hematopoietic stem cell make bone marrow transplantation clinically feasible, including its remarkable regenerative capacity, its ability to home to the marrow space following intravenous injection, and the ability of the stem cell to be cryopreserved. Transplantation of a single stem cell can replace the entire lymphohematopoietic system of an adult mouse. In humans, transplantation of a few percent of a donor's bone marrow volume regularly results in complete and sustained replacement of the recipient's entire lymphohematopoietic system, including all red cells, granulocytes, B and T lymphocytes, and platelets, as well as cells comprising the fixed macrophage population, including Kupffer cells of the liver, pulmonary alveolar macrophages, osteoclasts, Langerhans cells of the skin, and brain microglial cells. The ability of the hematopoietic stem cell to home to the marrow following intravenous injection is mediated, at least in part, by the interaction of specific cell molecules, termed *selectins*, on bone marrow endothelial cells with their unique ligands, termed *integrins*, on early hematopoietic cells. Human hematopoietic stem cells can survive freezing and thawing with little, if any, damage, making it possible to remove and store a portion of the patient's own bone marrow for later reinfusion following treatment of the patient with high-dose myelotoxic therapy.

CATEGORIES OF BONE MARROW TRANSPLANTATION

Bone marrow transplantation can be described according to the relationship between the patient and the donor and by the anatomic source of stem cells. In approximately 1% of cases, patients have identical twins who can serve as donors. Syngeneic donors represent the best source of stem cells; unlike allogeneic donors, there is no risk of graft-versus-host disease (GVHD) and, unlike use of autologous marrow, there is no risk that the stem cells are contaminated with tumor cells.

Allogeneic transplantation involves a donor and recipient who are not immunologically identical. Following allogeneic transplantation immune cells transplanted with the marrow or developing from it can react against the patient, causingGVHD. Alternatively, if the immunosuppressive preparative regimen used to treat the patient before transplant is inadequate, immunocompetent cells of the patient can cause graft rejection. The risks of these complications are greatly influenced by the degree of matching between donor and recipient for antigens encoded by genes of the major

histocompatibility complex.

The human leukocyte antigen (HLA) molecules are responsible for binding antigenic proteins and presenting them to T cells. The antigens presented by HLA molecules may derive from exogenous sources (e.g., during active infections) or may be endogenous proteins produced by the cell. If individuals are not matched for HLA, T cells from one individual will react strongly to the mismatched HLA, or "major antigens," of the second. Even if the individuals are HLA-matched, the T cells of the donor may react to differing endogenous, or "minor antigens," presented by the HLA of the recipient. Reactions to minor antigens tend to be less vigorous. The genes of major relevance to transplantation include HLA-A, -B, -C, and -D; they are closely linked and therefore tend to be inherited as haplotypes, with only rare crossovers between them. Thus, the odds that any one full sibling will match a patient are one in four, and the probability that the patient has an HLA-identical sibling is $1 - (0.75)^n$ where n equals the number of siblings.

With current techniques, the risk of graft rejection is 1 to 3%, and the risk of severe, life-threatening acuteGVHD is approximately 15% following transplantation betweenHLA-identical siblings. The incidence of graft rejection and GVHD increases progressively with the use of family member donors mismatched for one, two, or three antigens. While survival following a one-antigen mismatched transplant is not markedly altered, survival following two- or three-antigen mismatched transplants is significantly impaired, and such transplants should only be performed as part of clinical trials.

The formation of the National Marrow Donor Program has allowed for the identification ofHLA-matched unrelated donors for many patients. The genes encoding HLA antigens are highly polymorphic, and thus the odds of any two unrelated individuals being HLA-identical are extremely low, somewhat less than 1 in 10,000. However, by identifying and typing >3 million volunteer donors, HLA-matched donors now can found for approximately 50% of patients for whom a search is initiated. It takes, on average, 3 to 4 months to complete a search and schedule and initiate an unrelated donor transplant. Results so far suggest thatGVHD is somewhat increased and survival somewhat poorer with such donors than with HLA-matched siblings.

Autologous transplantation involves the removal and storage of the patient's own stem cells with subsequent reinfusion after the patient receives high-dose myeloablative therapy. Unlike allogeneic transplantation, there is no risk ofGVHD or graft rejection with autologous transplantation. On the other hand, autologous transplantation lacks a graft-versus-tumor effect, and the autologous stem cell product can be contaminated with tumor cells that could lead to relapse. A variety of techniques have been developed to "purge" autologous products of tumor cells. Some use antibodies directed at tumor-associated antigens plus complement, antibodies linked to toxins, or antibodies conjugated to immunomagnetic beads. In vitro incubation with certain chemotherapeutic agents such as 4-hydroperoxycylophosphamide and long-term culture of bone marrow has also been shown to diminish tumor cell numbers in stem cell products. Another technique is positive selection of stem cells using antibodies to CD34, with subsequent column adherence or flow techniques to select normal stem cells while leaving tumor cells behind. All these approaches can reduce the number of tumor cells from 1000- to 10,000-fold and are clinically feasible; however, no prospective randomized trials have yet shown that any of these approaches results in a decrease in relapse rates or

improvements in disease-free or overall survival.

Bone marrow aspirated from the posterior and anterior iliac crests has traditionally been the source of hematopoietic stem cells for transplantation. Typically, anywhere from 1.5 to 5´10^8nucleated marrow cells per kilogram are collected for allogeneic transplantation. Several recent studies have found improved survival in the settings of both matched sibling and unrelated transplantation by transplanting higher numbers of bone marrow cells.

Hematopoietic stem cells circulate in the peripheral blood but in very low concentrations. Following the administration of certain hematopoietic growth factors, including granulocyte colony stimulating factor (G-CSF) or granulocyte-macrophage colony stimulating factor (GM-CSF), and during recovery from intensive chemotherapy, the concentration of hematopoietic progenitor cells in blood, as measured either by colony forming units or expression of the CD34 antigen, increases markedly. This has made it possible to harvest adequate numbers of stem cells from the peripheral blood for transplantation. Donors are typically treated with 4 or 5 days of hematopoietic growth factor, following which stem cells are collected in one or two 4-h pheresis sessions. In the autologous setting, transplantation of>2.5 ´10^6 CD34 cells per kilogram, a number easily collected in most circumstances, leads to rapid and sustained engraftment in virtually all cases. Compared to the use of autologous marrow, use of peripheral blood stem cells results in more rapid hematopoietic recovery, with granulocytes recovering to 500/uL by day 12 and platelets recovering to 20,000/uL by day 14. While this more rapid recovery diminishes the morbidity of transplantation, no studies show an improvement in survival.

Hesitation in studying the use of peripheral blood stem cells for allogeneic transplantation was because peripheral blood stem cell products contain as much as one log more T cells than are contained in the typical marrow harvest; in animal models, the incidence ofGVHD is related to the number of T cells transplanted. Nonetheless, phase II and now randomized phase III trials have shown that the use of growth factor-mobilized peripheral blood stem cells fromHLA-matched family members leads to faster engraftment without an increase in acute GVHD. Chronic GVHD may be increased with peripheral blood stem cells, but in trials conducted so far, this has been more than balanced by reductions in relapse rates and nonrelapse mortality, with the use of peripheral blood stem cells resulting in improved overall survival.

Umbilical cord blood contains a high concentration of hematopoietic progenitor cells, allowing for its use as a source of stem cells for transplantation. Cord blood transplantation from family members has been explored in the setting where the immediate need for transplantation precludes waiting the 9 or so months generally required for the baby to mature to the point of donating marrow. Use of cord blood in such settings results in somewhat slower engraftment than seen with marrow but a low incidence ofGVHD, perhaps reflecting the low number of T cells in cord blood. More recently, several banks have been developed to harvest and store cord blood for possible transplantation to unrelated patients from material that would otherwise be discarded. A summary of the first 272 unrelated cord blood transplants, facilitated by the New York Blood Center, reported engraftment in approximately 90% of patients but at a slower pace than seen with a marrow. Significant GVHD was seen in 40% of patients.

The risk of graft failure was related to the dose of cord blood cells per kilogram infused. The low cell content of most cord blood collections has limited the use of this approach as a source of stem cells for adult patients.

THE TRANSPLANT PREPARATIVE REGIMEN

The treatment regimen administered to patients immediately preceding transplantation is designed to eradicate the patient's underlying disease and, in the setting of allogeneic transplantation, immunosuppress the patient adequately to prevent rejection of the transplanted marrow. The appropriate regimen, therefore, depends on the disease setting and source of marrow. For example, when transplantation is performed to treat severe combined immunodeficiency and the donor is a histocompatible sibling, no treatment is required because no host cells require eradication and the patient is already too immunoincompetent to reject the transplanted marrow. For aplastic anemia, there is no large population of cells to eradicate and high-dose cyclophosphamide plus antithymocyte globulin is sufficient to immunosuppress the patient adequately to accept the marrow graft. In the setting of thalassemia and sickle cell anemia, high-dose busulfan is frequently added to cyclophosphamide in order to eradicate the hyperplastic host hematopoiesis. A variety of different regimens have been developed to treat malignant diseases. Most of these regimens included agents that have high activity against the tumor in question at conventional doses and have myelosuppression as their predominant dose-limiting toxicity. Therefore, these regimens commonly include busulfan, cyclophosphamide, melphalan, thiotepa, carmustine, etoposide, and total-body irradiation in various combinations.

THE TRANSPLANT PROCEDURE

Marrow is usually collected from the donor's posterior and sometimes anterior iliac crests with the donor under general or spinal anesthesia. Typically, 10 to 15 mL/kg of marrow is aspirated, placed in heparinized media, and filtered through 0.3- and 0.2-mm screens to remove fat and bony spicules. The collected marrow may undergo further processing depending on the clinical situation, such as the removal of red cells to prevent hemolysis in ABO-incompatible transplants, the removal of donor T cells to preventGVHD, or attempts to remove possible contaminating tumor cells in autologous transplantation. Marrow donation is a safe procedure, with only very rare complications reported.

Peripheral blood stem cells are collected by leukopheresis after the donor has been treated with hematopoietic growth factors or, in the setting of autologous transplantation, sometimes after treatment with a combination of chemotherapy and growth factors. Stem cells for transplantation are generally infused through a large-bore central venous catheter. Such infusions are usually well tolerated, although occasionally patients develop fever, cough, or shortness of breath. These symptoms usually resolve with slowing of the infusion. When the stem cell product has been cryopreserved using dimethyl sulfoxide, patients more often experience short-lived nausea or vomiting due to the odor and taste of the cryoprotectant.

ENGRAFTMENT

Peripheral blood counts usually reach their nadir several days to a week posttransplant as a consequence of the preparative regimen, then cells produced by the transplanted stem cells begin to appear in the peripheral blood. The rate of recovery depends on the source of stem cells, the use of posttransplant growth factors, and the form ofGVHDprophylaxis employed. If marrow is the source of stem cells, recovery to 100 granulocytes per microliter occurs by day 16 and 500/uL by day 22. Use ofG-CSF-mobilized peripheral blood stem cells speeds the rate of recovery by approximately 1 week when compared to marrow. Use of myeloid growth factor (G-CSF orGM-CSF) posttransplant can further accelerate recovery by 3 to 5 days, while use of methotrexate to prevent GVHD delays engraftment by a similar period. Following allogeneic transplantation, engraftment can be documented using fluorescence in situ hybridization of sex chromosomes if donor and recipient are sex-mismatched,HLA-typing if HLA-mismatched, or restriction fragment length polymorphism analysis if sex- and HLA-matched.

COMPLICATIONS FOLLOWING BONE MARROW TRANSPLANT

EARLY DIRECT CHEMORADIOTOXICITIES

The transplant preparative regimens commonly used cause a spectrum of acute toxicities that vary according to the specific regimen but frequently result in nausea, vomiting, and mild skin erythema (Fig. 115-1). Regimens that include high-dose cyclophosphamide can result in hemorrhagic cystitis, which can usually be prevented by bladder irrigation or therapy with the sulfhydryl compound, mercaptoethanesulfonate (MESNA); rarely, acute hemorrhagic carditis is seen. Most preparative regimens will result in oral mucositis, which typically develops approximately 5 to 7 days posttransplant and often requires narcotic analgesia. Use of a patient-controlled analgesic pump provides the greatest patient satisfaction and results in a lower cumulative dose of narcotic. Patients begin losing their hair 5 to 6 days posttransplant and by 1 week are usually profoundly pancytopenic.

Approximately 10% of patients will develop venoocclusive disease of the liver, a syndrome resulting from direct cytotoxic injury to hepatic-venular and sinusoidal endothelium, with subsequent deposition of fibrin and the development of a local hypercoagulable state. This chain of events results in the clinical symptoms of tender hepatomegaly, ascites, jaundice, and fluid retention. These symptoms can develop any time during the first month posttransplant, with the peak incidence at day 16. The mortality of venoocclusive disease is approximately 30%, with progressive hepatic failure culminating in a terminal hepatorenal syndrome. Both thrombolytic and antithrombotic agents, such as tissue plasminogen activator, heparin, and prostaglandin E, have been studied as therapy, but none has proven of consistent major benefit in controlled trials and all have significant toxicity. Early studies with defibrotide, a polydeoxyribonucleotide, seem encouraging.

Although most pneumonias developing posttransplant are caused by infectious agents, in approximately 5% of patients a diffuse interstitial pneumonia will develop that is thought to be the result of direct toxicity of the preparative regimen. Bronchoalveolar lavage typically shows alveolar hemorrhage, and biopsies are typically characterized by diffuse alveolar damage, although some cases may have a more clearly interstitial

pattern. High-dose glucocorticoids are often used as treatment, although randomized trials testing their utility have not been reported.

LATE DIRECT CHEMORADIOTOXICITIES

Late complications of the preparative regimen include decreased growth velocity in children and delayed development of secondary sex characteristics. These complications can be partly ameliorated with the use of appropriate growth and sex hormone replacement. Most men become azoospermic, and most postpubertal women will develop ovarian failure, which should be treated. Thyroid dysfunction, usually well compensated, is sometimes seen. Cataracts develop in 10 to 20% of patients and are most common in patients treated with total-body irradiation and those who receive glucocorticoid therapy posttransplant for treatment ofGVHD. Aseptic necrosis of the femoral head is seen in 10% of patients and is particularly frequent in those receiving chronic glucocorticoid therapy.

GRAFT-VERSUS-HOST DISEASE

GVHDis the result of allogeneic T cells that were either transferred with the donor's stem cell inoculum or develop from it, reacting with antigenic targets on host cells. GVHD developing within the first 3 months posttransplant is termed *acute GVHD*, while GVHD developing or persisting beyond 3 months posttransplant is termed *chronic GVHD*. Acute GVHD most often first becomes apparent between 2 and 4 weeks posttransplant and is characterized by an erythematous maculopapular rash; persistent anorexia or diarrhea, or both; and by liver disease with increased serum levels of bilirubin, alanine and aspartate aminotransferase, and alkaline phosphatase. Since many conditions can mimic acute GVHD, diagnosis usually requires skin, liver, or endoscopic biopsy for confirmation. In all these organs, endothelial damage and lymphocytic infiltrates are seen. In skin, the epidermis and hair follicles are damaged; in liver, the small bile ducts show segmental disruption; and in intestines, destruction of the crypts and mucosal ulceration may be noted. A commonly used rating system for acute GVHD is shown inTable 115-1. Grade I acute GVHD is of little clinical significance, does not affect the likelihood of survival, and does not require treatment. In contrast, grades II to IV GVHD are associated with significant symptoms and a poorer probability of survival and require aggressive therapy. The incidence of acute GVHD is higher in recipients of stem cells from mismatched or unrelated donors, in older patients, and in patients unable to receive full doses of drugs used to prevent the disease.

One general approach to the prevention of GVHDis the administration of immunosuppressive drugs early after transplant. Combinations of methotrexate and either cyclosporine or tacrolimus are among the most effective and widely used regimens. Prednisone, anti-T cell antibodies, mycophenolate mofetil, and other immunosuppressive agents have also been or are being studied in various combinations. A second general approach to GVHD prevention is removal of T cells from the stem cell inoculum. While effective in preventing GVHD, T cell depletion is associated with an increased incidence of graft failure and of tumor recurrent posttransplant; as yet, little evidence suggests that this approach improves cure rates in any specific setting.

Despite prophylaxis, significant acuteGVHDwill develop in ~30% of recipients of stem cells from matched siblings and in as many as 60% of those receiving stem cells from unrelated donors. The disease is usually treated with glucocorticoids, anti-thymocyte globulin, or monoclonal antibodies targeted against T cells or T cell subsets.

Between 20 and 50% of patients surviving >6 months after allogeneic transplantation will develop chronicGVHD. The disease is more common in older patients, in recipients of mismatched or unrelated stem cells, and in those with a preceding episode of acute GVHD. The disease resembles an autoimmune disorder with malar rash, sicca syndrome, arthritis, obliterative bronchiolitis, and bile duct degeneration and cholestasis. Single-agent prednisone or cyclosporine is standard treatment at present, although trials of other agents, including thalidomide, are under way. In most patients, chronic GVHD resolves, but it may require 1 to 3 years of immunosuppressive treatment before these agents can be withdrawn without the disease recurring. Because patients with chronic GVHD are susceptible to significant infection, they should receive prophylactic trimethoprim-sulfamethoxazole, and all suspected infections should be investigated and treated aggressively.

GRAFT FAILURE

While complete and sustained engraftment are usually seen posttransplant, occasionally marrow function either does not return or, after a brief period of engraftment, is lost. Graft failure after autologous transplantation can be the result of inadequate numbers of stem cells being transplanted, damage during ex vivo treatment or storage, or exposure of the patient to myelotoxic agents posttransplant. Infections with cytomegalovirus (CMV) or human herpes virus type 6 have also been associated with loss of marrow function. Graft failure after allogeneic transplantation can also be due to immunologic rejection of the graft by immunocompetent host cells. Immunologically based graft rejection is more common following use of less immunosuppressive preparative regimens, in recipients of T cell-depleted stem cell products, and in patients receiving grafts fromHLA-mismatched donors.

Treatment of graft failure usually involves removing all potentially myelotoxic agents from the patient's regimen and attempting a short trial of myeloid growth factor. Persistence of lymphocytes of host origin in allogeneic transplant recipients with graft failure indicates immunologic rejection. Reinfusion of donor stem cells in such patients is usually unsuccessful unless preceded by a second immunosuppressive preparative regimen. Standard preparative regimens are generally tolerated poorly if administered within 100 days of a first transplant because of cumulative toxicities. However, use of regimens combining, for example, anti-CD3 antibodies with high-dose glucocorticoids have been successful in achieving engraftment in >50% of patients.

INFECTION

The general problem of infection in the immunocompromised host is discussed inChap. 136. Posttransplant patients, particularly recipients of allogeneic transplantation, require unique approaches. Early after transplantation, patients are profoundly neutropenic, and because the risk of bacterial infection is so great, most centers initiate antibiotic treatment once the granulocyte count falls to <500/uL. Fluconazole prophylaxis at a

dose of 200 to 400 mg/kg per day reduces the risk of candidal infections. Patients seropositive for herpes simplex should receive acyclovir prophylaxis. One approach to infection prophylaxis is shown inTable 115-2. Despite these prophylactic measures, most patients will develop fever and signs of infection posttransplant. The management of patients who become febrile despite bacterial and fungal prophylaxis is a difficult challenge and is guided by individual aspects of the patient and by the institution's experience.

Once patients engraft, the incidence of bacterial infection diminishes; however, patients, particularly allogeneic transplant recipients, remain at significant risk of infection. During the period from engraftment until about 3 months posttransplant, the most common causes of infection are gram-positive bacteria, fungi (particularly *Aspergillus*) and viruses includingCMV. CMV infection, which in the past was frequently seen and often fatal, can be prevented in seronegative patients by the use of seronegative blood products. The use of ganciclovir, either as prophylaxis beginning at the time of engraftment or initiated when CMV first reactivates as evidenced by development of antigenemia, can significantly reduce the risk of CMV disease in seropositive patients. Foscarnet is effective for some patients who develop CMV antigenemia or infection despite the use of ganciclovir or who cannot tolerate the drug.

Pneumocystis carinii pneumonia, once seen in 5 to 10% of patients, can be prevented by treating patients with oral trimethoprim-sulfamethoxazole for 1 week pretransplant and resuming the treatment once patients have engrafted.

The risk of infection diminishes considerably beyond 3 months after transplant unless chronicGVHDdevelops, requiring continuous immunosuppression. Most transplant centers recommend continuing trimethoprim-sulfamethoxazole prophylaxis while patients are receiving any immunosuppressive drugs and also recommend careful monitoring for lateCMVreactivation. In addition, most centers recommend prophylaxis against varicella zoster, using acyclovir for 1 year posttransplant.

TREATMENT OF SPECIFIC DISEASES USING BONE MARROW TRANSPLANTATION

NONMALIGNANT DISEASES

Immunodeficiency Disorders By replacing abnormal stem cells with cells from a normal donor, marrow transplantation can cure patients of a variety of immunodeficiency disorders including severe combined immunodeficiency, Wiskott-Aldrich syndrome, and Chediak-Higashi syndrome. The widest experience has been with severe combined immunodeficiency disease, where cure rates of 90% can be expected withHLA-identical donors and success rates of 50 to 70% have been reported using haplotype-mismatched parents as donors (Table 115-3).

Aplastic Anemia Transplantation from matched siblings after a preparative regimen of high-dose cyclophosphamide and antithymocyte globulin can cure up to 90% of patients younger than age 40 with severe aplastic anemia. Results in older patients and in recipients of mismatched family member or unrelated marrow are less favorable; therefore, a trial of immunosuppressive therapy is generally recommended for such

patients before considering transplantation. Transplantation is effective in all forms of aplastic anemia including, for example, the syndromes associated with paroxysmal nocturnal hemoglobinuria and Fanconi's anemia. Patients with Fanconi's anemia are abnormally sensitive to the toxic effects of alkylating agents and so less intensive preparative regimens must be used in their treatment (Chap. 109).

Hemoglobinopathies Marrow transplantation from anHLA-identical sibling following a preparative regimen of busulfan and cyclophosphamide can cure 70 to 90% of patients with thalassemia major. The best outcomes can be expected if patients are transplanted before they develop hepatomegaly or portal fibrosis and if they have been given adequate iron chelation therapy. Among such patients, the probabilities of 5-year survival and disease-free survival are 95 and 90%, respectively. Although prolonged survival can be achieved with aggressive chelation therapy, transplantation is the only curative treatment for thalassemia. Transplantation is being studied as a curative approach to patients with sickle cell anemia. Two-year survival and disease-free survival rates of 90 and 80%, respectively, have been reported following matched sibling transplantation. Decisions about patient selection and the timing of transplantation remain difficult, but transplantation seems to represent a reasonable option for younger patients who suffer repeated crises or other significant complications and who have not responded to other interventions (Chap 106).

Other Nonmalignant Diseases Theoretically, marrow transplantation should be able to cure any disease that results from an inborn error of the lymphohematopoietic system. Transplantation has been used successfully to treat congenital disorders of white blood cells such as Kostmann's syndrome, chronic granulomatous disease, and leukocyte adhesion deficiency. Congenital anemias such as Blackfan-Diamond anemia can also be cured with transplantation. Infantile malignant osteopetrosis is due to an inability of the osteoclast to resorb bone, and since osteoclasts derive from the marrow, transplantation can cure this rare inherited disorder.

Marrow transplantation has been used as treatment for a number of storage diseases caused by enzymatic deficiencies, such as Gaucher's disease, Hurler's syndrome, Hunter's syndrome, and infantile metachromatic leukodystrophy. Transplantation for these diseases has not been uniformly successful, but treatment early in the course of these diseases, before irreversible damage to extramedullary organs has occurred, increases the chance for success.

Transplantation is being explored as a treatment for severe acquired autoimmune disorders. These trials are based on studies demonstrating that transplantation can reverse autoimmune disorders in animal models and on the observation that occasional patients with coexisting autoimmune disorders and hematologic malignancies have been cured of both with transplantation.

MALIGNANT DISEASES

Acute Leukemia Allogeneic marrow transplantation cures 15 to 20% of patients who do not achieve complete response from induction chemotherapy for acute myeloid leukemia (AML) and is the only form of therapy that can cure such patients. Cure rates of 30 to 35% are seen when patients are transplanted in second remission or in first

relapse. The best results with allogeneic transplantation are achieved when applied during first remission, with disease-free survival rates averaging between 55 and 60%. Chemotherapy alone can cure a portion of AML patients, and so the relative merits of transplanting all patients during first remission versus only transplanting very high risk patients and those who relapse continue to be discussed. Autologous transplantation is also able to cure a portion of patients with AML. The rates of disease recurrence with autologous transplantation are higher than seen after allogeneic transplantation, and cure rates are generally somewhat less.

Similar to patients withAML, adults with acute lymphoblastic leukemia who do not achieve a complete response to induction chemotherapy can be cured in 15 to 20% of cases with immediate marrow transplantation. Cure rates improve to 30 to 50% in second remission, and therefore transplantation can be recommended for adults who have persistent disease after induction chemotherapy or who have subsequently relapsed. Transplantation in first remission results in cure rates around 55%. While transplantation appears to offer a clear advantage over chemotherapy for patients with high-risk disease, such as those with Philadelphia chromosome-positive disease, debate continues about whether adults with standard-risk disease would be transplanted in first remission or whether transplantation should be reserved until relapse. Autologous transplantation is associated with a higher relapse rate but a somewhat lower risk of nonrelapse mortality when compared to allogeneic transplantation. On balance, most experts recommend use of allogeneic stem cells if an appropriate donor is available.

Chronic Leukemia Allogeneic marrow transplantation is the only therapy shown to cure a substantial portion of patients with chronic myeloid leukemia. Five-year disease-free survival rates are 60 to 70% for patients transplanted during chronic phase, 30 to 40% for patients transplanted during accelerated phase, and 15 to 20% for patients transplanted in blast crisis. Time from diagnosis to transplantation influences outcome, with best results obtained among patients transplanted within 1 year of diagnosis. Use of unrelated donors results in moreGVHD and slightly worse survival than seen with matched siblings, although, at some large centers, 3-year disease-free survival rates of 70% have been reported. Autologous transplantation is being studied; however, few data suggest that this approach has curative potential in this disease. Given the excellent results obtained with matched sibling transplantation, most experts recommend early transplantation for younger patients with matched siblings. For older patients or those without matched siblings, it is not unreasonable to consider a trial of an interferon a-containing regimen to see if a major cytogenetic response can be achieved before making a decision about transplantation (Chap. 111).

Allogeneic transplantation has been used to only a limited extent for chronic lymphocytic leukemia, in large part because of the chronic nature of the disease and because of the age profile of patients. With allogeneic transplantation, complete remissions have been achieved in the majority of patients so far reported, with disease-free survival rates of approximately 50% at 3 years. However, treatment-related mortality has been substantial, and further follow-up is needed. There is even less experience with autologous transplantation in this disorder.

Myelodysplasia Between 40 and 50% of patients with myelodysplasia appear to be

cured with allogeneic marrow transplantation. Results are better among younger patients and those with less advanced disease. However, some patients with myelodysplasia can live for extended periods without intervention, and so transplantation is generally recommended only for patients with disease categorized as intermediate risk I or greater according to the International Prognostic Scoring System (Chap. 109).

Lymphoma Patients with disseminated intermediate- or high-grade non-Hodgkin's lymphoma who have not been cured by first-line chemotherapy and are transplanted in first relapse or second remission can still be cured in 40 to 50% of cases. This represents a clear advantage over results obtained with salvage chemotherapy. It is unsettled whether patients with high-risk disease benefit from transplantation in first remission. Most experts favor the use of autologous rather than allogeneic transplantation for patients with non-Hodgkin's lymphoma, because fewer complications occur with this approach and survival appears equivalent. The role of transplantation in patients with indolent non-Hodgkin's lymphoma is less well defined. Long-term remissions can be obtained in many patients with acceptable toxicity and results with transplantation in patients with recurrent disease generally appear better than one would expect with conventional-dose chemotherapy. However, late relapses are seen after transplantation, and no randomized study has confirmed its superiority.

The role of transplantation in Hodgkin's disease is similar to that in non-Hodgkin's lymphoma. With transplantation, 5-year disease-free survival ranges from 20 to 30% in patients who never achieve a first remission with standard chemotherapy and up to 60% for those transplanted in second remission. Transplantation has no defined role in first remission in Hodgkin's disease.

Myeloma Patients with myeloma who have progressed on first-line therapy can sometimes benefit from allogeneic or autologous transplantation. Autologous transplantation has been studied as part of the initial therapy of patients, and in randomized trials, both disease-free survival as well as overall survival were improved with this approach.

Solid Tumors Among women with metastatic breast cancer, between 15 and 20% disease-free survival rates at 3 years have been reported, with better results seen in younger patients who have responded completely to standard-dose therapy before undergoing transplantation. Randomized trials have not shown superior survival for patients treated for metastatic disease with high-dose chemotherapy plus stem cell support. Randomized trials evaluating transplantation as treatment for primary breast cancer are being conducted, but final results are not yet available.

Patients with testicular cancer who have failed first-line chemotherapy have been treated with autologous transplantation. Approximately 10 to 20% of such patients apparently have been cured with this approach.

The use of high-dose chemotherapy with autologous stem cell support is being studied for several other solid tumors, including ovarian cancer, small-cell lung cancer, neuroblastoma, and pediatric sarcomas. As in most other settings, the best results have been obtained in patients with limited amounts of disease and where the remaining

tumor retains sensitivity to conventional-dose chemotherapy. Few randomized trials of transplantation in these diseases have been completed.

Posttransplant Relapse Patients who relapse following autologous transplantation sometimes respond to further chemotherapy, particularly if the remission following transplantation was long. More options are available for patients who relapse following allogeneic transplantation. Of particular interest are the response rates seen with infusion of unirradiated donor lymphocytes. Complete responses in as many as 75% of patients with chronic myeloid leukemia, 40% in myelodysplasia, 25% inAML, and 15% in myeloma have been reported. Major complications of donor lymphocyte infusions include transient myelosuppression and the development ofGVHD. These complications appear to be dependent on the number of donor lymphocytes infused. The impressive responses seen with donor lymphocyte infusions in some patients has encouraged investigation into the use of "nonablative" transplant regimens as treatment for various malignancies. In this approach, preparative regimens and posttransplant immunosuppression are selected that allow for engraftment without regard to their direct antitumor activities. The antitumor effects are the result of a graft-versus-tumor effect arising from the transplanted stem cells or subsequent infusion of donor lymphocytes. While engraftment can be reliably achieved with this approach, with little toxicity, and complete responses are seen, neither the rate of complete responses nor their duration have yet been entirely determined for any specific disease category.

(Bibliography omitted in Palm version)

Back to Table of Contents

SECTION 3 -DISORDERS OF HEMOSTASIS

116. DISORDERS OF THE PLATELET AND VESSEL WALL - *Robert I. Handin*

Patients with platelet or vessel wall disorders usually bleed into superficial sites such as the skin, mucous membranes, or genitourinary or gastrointestinal tract. Bleeding begins immediately after trauma and either responds to simple measures such as pressure and packing or requires systemic therapy with glucocorticoids, desmopressin [1-desamino-8-D-arginine vasopressin (DDAVP)], plasma fractions, or platelet concentrates. The most common platelet/vessel wall disorders are (1) various forms of thrombocytopenia, (2) von Willebrand's disease (vWD), and (3) drug-induced platelet dysfunction. This chapter reviews the diagnosis and treatment of quantitative and qualitative platelet disorders as well as vessel wall defects that cause bleeding.*For further discussion of the physiology of normal hemostasis and the cardinal manifestations of bleeding arising from hemostatic disorders, see Chap. 62.*

PLATELET DISORDERS

Platelets arise from the fragmentation of megakaryocytes, which are very large, polyploid bone marrow cells produced by the process of endomitosis. They undergo from three to five cycles of chromosomal duplication without cytoplasmic division. After leaving the marrow space, about one-third of the platelets are sequestered in the spleen, while the other two-thirds circulate for 7 to 10 days. Normally, only a small fraction of the platelet mass is consumed in the process of hemostasis, so most platelets circulate until they become senescent and are removed by phagocytic cells. The normal blood platelet count is 150,000 to 450,000/ul. A decrease in platelet count stimulates an increase in the number, size, and ploidy of megakaryocytes, releasing additional platelets into the circulation. This process is regulated by thrombopoietin (TPO) binding to its megakaryocyte receptor, a proto-oncogene c-mpl. TPO (c-mpl ligand) is secreted continuously at a low level and binds tightly to circulating platelets. A reduction in platelet count increases the level of free TPO and thereby stimulates megakaryocyte and platelet production.

The platelet count varies during the menstrual cycle, rising following ovulation and falling at the onset of menses. It is also influenced by the patient's nutritional state and can be decreased in severe iron, folic acid, or vitamin B_{12}deficiency. Platelets are *acute-phase reactants*, and patients with systemic inflammation, tumors, bleeding, and mild iron deficiency may have an increased platelet count, a benign condition called *secondary* or *reactive thrombocytosis*. The cytokines interleukin (IL)-3, IL-6, and IL-11 may stimulate platelet production in acute inflammation. In contrast, the increase in platelet count that is characteristic of the myeloproliferative disorders such as polycythemia vera, chronic myelogenous leukemia, myeloid metaplasia, and essential thrombocytosis can cause either severe bleeding or thrombosis. In these patients, unregulated platelet production is secondary to a clonal stem cell abnormality affecting all the bone marrow progenitors.

THROMBOCYTOPENIA

Thrombocytopenia is caused by one of three mechanisms -- decreased bone marrow

production, increased splenic sequestration, or accelerated destruction of platelets. In order to determine the etiology of thrombocytopenia, each patient should have a careful examination of the peripheral blood film, an assessment of marrow morphology by examination of an aspirate or biopsy, and an estimate of splenic size by bedside palpation supplemented, if necessary, by ultrasonography or computed tomographic (CT) scan. Occasional patients have "pseudothrombocytopenia," a benign condition in which platelets agglutinate or adhere to leukocytes when blood is collected with EDTA as anticoagulant. This is a laboratory artifact, and the actual platelet count in vivo is normal. A scheme for classifying patients with thrombocytopenia based on these clinical observations and laboratory tests is outlined inFig. 116-1.

Impaired Production Disorders that injure stem cells or prevent their proliferation frequently cause thrombocytopenia. They usually affect multiple hematopoietic cell lines so that thrombocytopenia is accompanied by varying degrees of anemia and leukopenia. Diagnosis of a platelet production defect is readily established by examination of a bone marrow aspirate or biopsy, which should show a reduced number of megakaryocytes. The most common causes of decreased platelet production are marrow aplasia, fibrosis, or infiltration with malignant cells, all of which produce highly characteristic marrow abnormalities. Occasionally, thrombocytopenia is the presenting laboratory abnormality in these disorders. Cytotoxic drugs impair megakaryocyte proliferation and maturation and frequently cause thrombocytopenia. Rare marrow disorders such as congenital amegakaryocytic hypoplasia and *t*hrombocytopenia with *a*bsent *r*adii (TAR syndrome), produce a selective decrease in megakaryocyte production.

Splenic Sequestration Since one-third of the platelet mass is normally sequestered in the spleen, splenectomy will increase the platelet count by 30%. Postsplenectomy thrombocytosis is a benign self-limited condition that does not require specific therapy. In contrast, when the spleen enlarges, the fraction of sequestered platelets increases, lowering the platelet count. The most common causes of splenomegaly are portal hypertension secondary to liver disease and splenic infiltration with tumor cells in myeloproliferative or lymphoproliferative disorders (Chap. 63). Isolated splenomegaly is rare, and in most patients it is accompanied by other clinical manifestations of an underlying disease. Many patients with leukemia, lymphoma, or a myeloproliferative syndrome have both marrow infiltration and splenomegaly and develop thrombocytopenia from a combination of impaired marrow production and splenic sequestration of platelets.

Accelerated Destruction Abnormal vessels, fibrin thrombi, and intravascular prostheses can all shorten platelet survival and cause *nonimmunologic thrombocytopenia*. Thrombocytopenia is common in patients with vasculitis, the hemolytic uremic syndrome (HUS), thrombotic thrombocytopenic purpura (TTP), or as a manifestation of disseminated intravascular coagulation (DIC). In addition, platelets coated with antibody, immune complexes, or complement are rapidly cleared by mononuclear phagocytes in the spleen or other tissues, inducing *immunologic thrombocytopenia*. The most common causes of immunologic thrombocytopenia are viral or bacterial infections, drugs, and a chronic autoimmune disorder referred to as *idiopathic thrombocytopenic purpura* (ITP). Patients with immunologic thrombocytopenia do not usually have splenomegaly and have an increased number of bone marrow

megakaryocytes.

DRUG-INDUCED THROMBOCYTOPENIA

Many common drugs can cause thrombocytopenia (Table 116-1). Cancer chemotherapeutic agents may depress megakaryocyte production. Ingestion of large quantities of alcohol has a marrow-depressing effect leading to transient thrombocytopenia, particularly in binge drinkers. Thiazide diuretics, used to treat hypertension or congestive heart failure, impair megakaryocyte production and can produce mild thrombocytopenia (50,000 to 100,000/uL), which may persist for several months after the drug is discontinued.

Most drugs induce thrombocytopenia by eliciting an immune response in which the platelet is an innocent bystander. The platelet is damaged by complement activation following the formation of drug-antibody complexes. Current laboratory tests can identify the causative agent in 10% of patients with clinical evidence of drug-induced thrombocytopenia. The best proof of a drug-induced etiology is a prompt rise in the platelet count when the suspected drug is discontinued. Patients with drug-induced platelet destruction may also have a secondary increase in megakaryocyte number without other marrow abnormalities.

Although most patients recover within 7 to 10 days and do not require therapy, occasional patients with platelet counts <10,000 to 20,000/uL have severe hemorrhage and may require temporary support with glucocorticoids, plasmapheresis, or platelet transfusions while waiting for the platelet count to rise. A patient who has recovered from drug-induced immunologic thrombocytopenia should be instructed to avoid the offending drug in the future, since only minute amounts of drug are needed to set up subsequent immune reactions. Certain drugs that are cleared from body storage depots quite slowly, such as phenytoin, may induce prolonged thrombocytopenia.

Heparin is a common cause of thrombocytopenia in hospitalized patients. Between 10 and 15% of patients receiving therapeutic doses of heparin develop thrombocytopenia and, occasionally, may have severe bleeding or intravascular platelet aggregation and paradoxical thrombosis. Heparin-induced thrombosis, sometimes called the "white clot syndrome," can be fatal unless recognized promptly. Most cases of heparin thrombocytopenia are due to drug-antibody binding to platelets; some are secondary to direct platelet agglutination by heparin. The offending antigen is a complex formed between heparin and the platelet-derived heparin neutralizing protein, platelet factor 4. Prompt cessation of heparin will reverse both thrombocytopenia and heparin-induced thrombosis. Low-molecular-weight heparin products have reduced the incidence of heparin-induced thrombocytopenia. They are effective antithrombotic agents (Chap. 118) and are less immunogenic. Unfortunately, 80 to 90% of the antibodies generated against conventional heparins cross-react with low-molecular-weight heparins, so only a minority of patients with preformed antibody can be treated with this product.

IDIOPATHIC THROMBOCYTOPENIC PURPURA

The immunologic thrombocytopenias can be classified on the basis of the pathologic mechanism, the inciting agent, or the duration of the illness. The explosive onset of

severe thrombocytopenia following recovery from a viral exanthem or upper respiratory illness (*acuteITP*) is common in children and accounts for 90% of the pediatric cases of immunologic thrombocytopenia. Of these patients, 60% recover in 4 to 6 weeks and>90% recover within 3 to 6 months. Transient immunologic thrombocytopenia also complicates some cases of infectious mononucleosis, acute toxoplasmosis, or cytomegalovirus infection and can be part of the prodromal phase of viral hepatitis and initial infection with HIV. Acute ITP is rare in adults and accounts for<10% of postpubertal patients with immune thrombocytopenia. Acute ITP is caused by immune complexes containing viral antigens that bind to platelet Fc receptors or by antibodies produced against viral antigens that cross-react with the platelet. In addition to these viral disorders, the differential diagnosis includes atypical presentations of aplastic anemia, acute leukemias, or metastatic tumor. A bone marrow examination is essential to exclude these disorders, which can occasionally mimic acute ITP.

Most adults present with a more indolent form of thrombocytopenia that may persist for many years and is referred to as *chronicITP*. Women age 20 to 40 are afflicted most commonly and outnumber men by a ratio of 3:1. They may present with an abrupt fall in platelet count and bleeding similar to patients with acute ITP. More often they have a prior history of easy bruising or menometrorrhagia. These patients have an autoimmune disorder with antibodies directed against target antigens on the glycoprotein IIb-IIIa or glycoprotein Ib-IX complex (Fig. 62-2). Although most antibodies function as opsonins and accelerate platelet clearance by phagocytic cells, occasional antibodies bind to epitopes on critical regions of these glycoproteins and impair platelet function. Platelet-associated IgG can be measured but specificity is a problem. High "background" level of IgG on normal platelets and elevations in plasma immunoglobulin levels or in circulating immune complexes will nonspecifically increase platelet-associated IgG. Few clinical situations require platelet-associated IgG testing.

A low platelet count may be the initial manifestation of systemic lupus erythematosus (SLE) or the first sign of a primary hematologic disorder. Thus, patients with chronicITPshould have a bone marrow examination and an antinuclear antibody determination. In addition, patients with hepatic or splenic enlargement, lymphadenopathy, or atypical lymphocytes should have serologic studies for hepatitis viruses, cytomegalovirus, Epstein-Barr virus, toxoplasma, and HIV. HIV infection is a common cause of immunologic thrombocytopenia. Thrombocytopenia can be the initial symptom of HIV infection or a complication of fully developed clinical AIDS.

TREATMENT

Treatment of patients withITP must take into account the age of the patient, the severity of the illness, and the anticipated natural history. Although adults have a higher incidence of intracranial bleeding than children, specific therapy may not be necessary unless the platelet count is <20,000/uL or there is extensive bleeding. Hemorrhage in patients with either acute or chronic ITP can usually be controlled with glucocorticoids but, in rare cases, may require temporary phagocytic blockade with intravenous immunoglobulin (IVIG) or anti-RhD (WinRho). Although antibody preparations are effective, they are expensive and should be reserved for patients with severe thrombocytopenia and clinical bleeding who are refractory to other measures. Emergency splenectomy is usually reserved for patients with acute or chronic ITP who

are desperately ill and have not responded to any medical measures. The treatment of symptomatic thrombocytopenia in patients with HIV infection is more complex because the administration of glucocorticoids or splenectomy may increase susceptibility to opportunistic infections. Splenectomy has been effective in the course of HIV before the onset of symptomatic AIDS. Treatment with zidovudine (AZT) and other antiviral agents that reduce viral load can improve the platelet count in patients with HIV-induced thrombocytopenia.

Symptomatic patients with chronicITP are usually placed on prednisone, 60 mg/d for 4 to 6 weeks. The drug is then decreased slowly over another few weeks. About 50% of patients with chronic ITP will normalize their platelet count on high doses of prednisone. However, the majority will have a fall in platelet count following steroid withdrawal. Patients with chronic ITP who fail to maintain a normal platelet count after a course of prednisone are eligible for elective splenectomy. These steroid-responsive but steroid-dependent patients are very likely to respond to splenectomy, and 70% will have a normal platelet count within 1 week after surgery. Some patients who do not respond to glucocorticoids may still respond to splenectomy. Occasionally, patients may fail to respond to splenectomy because of the failure to remove an accessory spleen. In other patients, a small, inactive accessory spleen may grow or new splenic foci may develop from splenic cells shed at the time of surgery and cause the late onset of thrombocytopenia. In either case, the presence of splenic tissue can be diagnosed by examination of the blood smear for Howell-Jolly bodies that appear in the red cells of asplenic individuals. Persistent splenic tissue can be confirmed by a radionuclide scan.

Patients still thrombocytopenic after splenectomy or who relapse months to years after initial therapy have received a variety of immunosuppressive drugs including azathioprine, cyclophosphamide, vincristine, vinblastine, and cyclosporine. Danazol has also been used with some success. Although each of these drugs may be beneficial, they have serious side effects and should be used judiciously.IVIG and anti-RhD are only transiently effective and expensive. IVIG can cause meningismus and headache, and some lots have carried hepatitis C virus. Anti-RhD can cause hemolysis. These drugs should be used to raise the platelet count temporarily and to support patients before surgery or labor and delivery; they are not substitutes for splenectomy. If a patient is not bleeding and maintains a platelet count >20,000/uL, consideration should be given to withholding therapy. Patients with severe chronic thrombocytopenia may live with their disease for two or three decades.

FUNCTIONAL PLATELET DISORDERS

As described inChap. 62, normal hemostasis requires three critical platelet reactions -- adhesion, aggregation, and granule release. Clinical bleeding can result from a failure of any of these important functions. Table 116-2 lists the major functional platelet disorders. Table 116-3 lists methods to assess platelet function.

Von Willebrand's Disease vWD is the most common inherited bleeding disorder, occurring in 1 in 800 to 1000 individuals. The von Willebrand factor (vWF) is a heterogeneous multimeric plasma glycoprotein with two major functions: (1) It facilitates platelet adhesion under conditions of high shear stress by linking platelet membrane receptors to vascular subendothelium; and (2) it serves as the plasma carrier for factor

VIII, the antihemophilic factor, a critical blood coagulation protein. Discrete domains in each vWF subunit mediate each of these important functions. The normal plasma vWF level is 10 mg/L. The vWF activity is distributed among a series of plasma multimers with estimated molecular weights ranging from 400,000 to >20 million. A single large vWF precursor subunit is synthesized in endothelial cells and megakaryocytes, where it is cleaved and assembled into the disulfide-linked multimers present in plasma, platelets, and vascular subendothelium. A modest reduction in plasma vWF concentration or a selective loss in the high-molecular-weight multimers decreases platelet adhesion and causes clinical bleeding.

AlthoughvWD is heterogeneous, certain clinical features are common to all the syndromes. With one exception (type III disease), all forms are inherited as autosomal dominant traits, and affected patients are heterozygous with one normal and one abnormalvWFallele. In mild cases, bleeding occurs only after surgery or trauma. More severely affected patients have spontaneous epistaxis or oral mucosal, gastrointestinal, or genitourinary bleeding. The laboratory findings are variable. The most diagnostic pattern is the combination of (1) a prolonged bleeding time, (2) a reduction in plasma vWF concentration, (3) a parallel reduction in biologic activity as measured with the ristocetin cofactor assay, and (4) reduced factor VIII activity. The variability in laboratory tests is related to both the heterogeneous nature of the defects in vWD and the fact that plasma levels are influenced by ABO blood group type, central nervous system disorders, systemic inflammation, and pregnancy. Since vWD is an autosomal dominant disorder, some vWF is produced by the remaining normal allele. Thus patients with mild defects may have laboratory values that fluctuate over time and may occasionally be within the normal range.

There are three major types ofvWD. Their mode of inheritance and laboratory findings are shown inFig. 116-2. Patients with *type I disease*, the most common abnormality, have a mild to moderate decrease in plasmavWF. In the milder cases, although hemostasis is impaired, the vWF level is just below normal (50% activity, or 5 mg/L). In type I disease, vWF antigen, factor VIII activity, and ristocetin cofactor activity are decreased with a normal spectrum of multimers detected by sodium dodecyl sulfate (SDS)-agarose gel electrophoresis.

The variant forms ofvWD(*type II disease*) are much less common and characterized by normal or near-normal levels of a dysfunctional protein. Patients with the *type IIa variant* of vWD have a deficiency in the high- and medium-molecular-weight forms ofvWFmultimer detected bySDS-agarose electrophoresis. This is due either to an inability to secrete the high-molecular-weight vWF multimers or to proteolysis of the multimers soon after they leave the endothelial cell and enter the circulation. Mutations in a localized region of the vWF A-2 domain have been identified in families with type IIa vWD (Fig. 116-3). The quantity of vWF antigen and the amount of associated factor VIII are usually normal. In the *type IIb variant*, high-molecular-weight multimers are also decreased; however, the decrease is due to the inappropriate binding of vWF to platelets. Intravascular platelet aggregates form that are rapidly cleared from the circulation, causing mild, variable thrombocytopenia. Mutations in a disulfide-bonded loop in the A-1 domain that binds to glycoprotein Ib-IX are the cause of the type IIb defect (Fig. 116-3). A few patients have a platelet membrane disorder that mimics type IIb vWD -- *platelet-type vWD*. It is due to mutations in the portion of glycoprotein Ib-IX

that interacts with vWF. Levels of total vWF antigen and factor VIII are normal.

Approximately 1 in 1 million individuals has a very severe form ofvWD that is phenotypically recessive (*type III disease*). Type III patients are usually the offspring of two parents (usually asymptomatic) with mild type I disease. Type III patients may inherit a different abnormality from each parent (a doubly heterozygous or compound heterozygous state) or be homozygous for a single defect. Type III patients have severe mucosal bleeding and no detectablevWFantigen or activity and, like patients with mild hemophilia, may have sufficiently low factor VIII that they have occasional hemarthroses. Major deletions in the vWF gene have been found in some type III families. Families with nonsense mutations and the combination of a deleted and nonsense mutant allele have also been described.

Type IIN disease is due to a defect in the factor VIII binding site ofvWF. Patients resemble those with mild hemophilia and have low levels of factor VIII. The presence of disease in both males and females in a family is a clue to the role of vWF in this disease.

TREATMENT

There are two therapeutic options. Factor VIII concentrates retain high-molecular-weightvWFmultimers (Humate-P, Alfanate), are highly purified and heat-treated to destroy HIV, and are appropriate treatments for all the inherited forms ofvWD. During surgery or after major trauma, patients should receive factor VIII concentrates twice daily for 2 or 3 days to assure optimal hemostasis. Minor bleeding episodes such as prolonged epistaxis or severe menorrhagia may respond to a single infusion. Recurrent menorrhagia, a major problem for women with severe vWD, can be treated effectively with oral contraceptive agents that suppress menses.

A second therapeutic option, which avoids the use of plasma, is the use ofDDAVP or desmopressin, a vasopressin analogue that has minimal blood pressure-elevating and fluid-retaining properties and raises the plasmavWFlevel in both normal individuals and patients with mildvWD. Patients with type I disease are the best candidates for DDAVP therapy. However, they must be tested for an adequate response before anticipated surgery, and vWF levels must be monitored closely during therapy, since the patient may develop tachyphylaxis when therapy is continued for more than 48 h. DDAVP should not be given to patients with variant forms of vWD without prior testing, since it may not improve multimer pattern or hemostasis in type IIa patients and may actually worsen the defect by depleting high-molecular-weight multimers, inducing intravascular platelet aggregation, and lowering the platelet count in type IIb patients. It is ineffective therapy for the severe (type III) form of vWD.

Acquired vWD Although most cases of vWDare inherited, acquired vWD may be caused by antibodies that inhibitvWFfunction or by lymphoid or other tumors that selectively adsorb vWF multimers onto their surfaces. Anti-vWF antibodies have developed in patients with severe vWD following multiple transfusions, as well as in patients with autoimmune and lymphoproliferative disorders. Adsorption of vWF to tumor surfaces has been documented in patients with Waldenstrom's macroglobulinemia and Wilms' tumor and inferred in other patients with lymphoma. Treatment of acquired vWD should

focus on the underlying disease, since plasma derivatives andDDAVPare often not effective and the disorder can be fatal.

Platelet Membrane Defects Receptors that modulate platelet adhesion and aggregation are located on the two major platelet surface glycoproteins.vWFfacilitates platelet adhesion by binding to glycoprotein Ib-IX, while fibrinogen links platelets into aggregates via sites on the glycoprotein IIb-IIIa complex. Two rare platelet defects are characterized by a loss of or a defect in these glycoprotein receptors. Patients with the *Bernard-Soulier syndrome* have markedly reduced platelet adhesion and cannot bind vWF to their platelets due to deficiency or dysfunction of the glycoprotein Ib-IX complex. They also have reduced levels of several other membrane proteins, mild thrombocytopenia, and extremely large, lymphocytoid platelets. Platelets from patients with *Glanzmann's disease* or *thrombasthenia* are deficient or defective in the glycoprotein IIb-IIIa complex. Their platelets do not bind fibrinogen and cannot form aggregates, although the platelets undergo shape change and secretion and are of normal size.

Both these disorders are autosomal recessive traits and are characterized by markedly impaired hemostasis and recurrent episodes of severe mucosal hemorrhage. Bernard-Soulier platelets react normally to all stimuli except ristocetin. In contrast, thrombasthenic platelets adhere normally and will agglutinate with ristocetin but will not aggregate with any of the agonists that require fibrinogen binding, such as adenosine diphosphate (ADP), thrombin, or epinephrine.

The only effective therapy for hemorrhagic episodes in these two disorders is transfusion with normal platelets. Alloimmunization will eventually limit the life span of infused platelets. In addition, a few patients have developed inhibitor antibodies with specificity for the missing protein. These antibodies bind to the protein that is expressed on the transfused normal platelets and impair their function.

Platelet Release Defects The most common mild bleeding disorders arise from the ingestion of aspirin and other nonsteroidal anti-inflammatory drugs (NSAIDs) that inhibit platelet production of thromboxane A_2, an important mediator of platelet secretion and aggregation (Figs. 62-3 and62-4). These drugs inhibit cyclooxygenase, which converts arachidonic acid to a labile endoperoxide intermediate that is critical for thromboxane formation. Aspirin is the most potent agent, since it irreversibly acetylates the platelet enzyme so that a single dose impairs hemostasis for 5 to 7 days. The other agents are competitive and reversible inhibitors with more transient effects. Blocking thromboxane A_2synthesis partially inhibits platelet release and aggregation with weak agonists, such asADP and epinephrine, and produces a mild hemostatic defect. The administration of high doses of certain antibiotics, particularly penicillin, can coat the platelet surface, block platelet release, and impair hemostasis.

Patients generally have minimal symptoms such as easy bruising, and bleeding is usually confined to the skin. Occasional patients will have prolonged oozing after surgery, particularly with procedures involving mucous membranes such as periodontal, oral, or reconstructive plastic surgery. The antiplatelet effect of drugs such as aspirin is more dramatic when they are administered to patients with underlying defects such asvWD or hemophilia. Patients with drug-induced cyclooxygenase deficiency often have

a mildly prolonged bleeding time, and their platelets fail to aggregate when incubated with arachidonic acid, epinephrine, or low doses ofADP. Patients who have taken aspirin should be treated as if they have a mild hemostatic defect for the next 5 to 7 days. Platelet responses to collagen and thrombin are impaired at low doses but normal at higher doses. Symptomatic patients should be encouraged to use drugs such as acetaminophen that do not impair platelet function. Although most cases of cyclooxygenase deficiency are drug-induced, occasional patients have inherited disorders in platelet cyclooxygenase activity that impair thromboxane production or receptor level defects that prevent platelets from responding to thromboxane A_2.

Of the metabolic disorders that can perturb hemostasis, uremic platelet dysfunction is clinically the most important. The mechanism by which uremia impairs platelet function is not well understood, and retention of phenolic and guanidinosuccinic acids, excess prostacyclin production, or impairedvWF-platelet interactions have all been implicated. The degree of uremia correlates with bleeding symptoms and anemia. Bleeding can usually be reversed by dialysis and often improves after red cell transfusion or treatment with erythropoietin. In addition, factor VIII concentrate orDDAVP, both of which raise plasma vWF levels, can also improve hemostasis. Conjugated estrogens improve hemostasis and can be used as long-term therapy.

Storage Pool Defects Platelet granules have considerable amounts of adenine nucleotides, calcium, and adhesive glycoproteins such as thrombospondin, fibronectin, andvWF, all of which promote platelet adhesion and aggregation. Patients with defective platelet granules have a mild bleeding disorder. Platelet storage pool defects may be inherited as an isolated disorder or be part of systemic granule packaging defects such as oculocutaneous albinism or the Hermansky-Pudlak or Chediak-Higashi syndromes. Clinically, these patients cannot be distinguished from those with other functional platelet disorders, since they all have easy bruising, mucosal bleeding, and a prolonged bleeding time. They can be differentiated from patients with the cyclooxygenase defects because their platelets will usually aggregate in response to arachidonic acid. In addition, their platelets have decreased levels of specific granule constituents such asADP and serotonin and abnormalities in granule morphology that are best visualized by electron microscopy.

Occasionally, patients with acute or chronic leukemia or one of the myeloproliferative disorders develop an acquired storage pool disorder due to dysplastic megakaryocyte development. In addition, patients with liver disease and some patients withSLE or other immune complex-mediated disorders may have circulating platelets that have degranulated prematurely. Platelet degranulation and a transient storage pool disorder may occur after prolonged cardiopulmonary bypass. Fortunately, most patients with storage pool defects have only mildly impaired hemostasis. They can be treated with platelet transfusions. Occasional patients have responded toDDAVP.

VESSEL WALL DISORDERS

Bleeding from vascular disorders (nonthrombocytopenic purpura) is usually mild and confined to the skin and mucous membranes. The pathogenesis of bleeding is poorly defined in many of the syndromes, and classic tests of hemostasis, including the bleeding time and tests of platelet function, are usually normal. Vascular purpura arises

from damage to capillary endothelium, abnormalities in the vascular subendothelial matrix or extravascular connective tissues that support blood vessels, or from the formation of abnormal blood vessels. Several idiopathic disorders involve the vessel wall and can cause more severe bleeding and organ dysfunction.

THROMBOTIC THROMBOCYTOPENIC PURPURA

TTPis a fulminant, often lethal disorder that may be initiated by endothelial injury and subsequent release ofvWF and other procoagulant materials from the endothelial cell. Causes include pregnancy, metastatic cancer, mitomycin C, high-dose chemotherapy, HIV infection, and certain drugs, such as the antiplatelet angent ticlopidine. Characteristic findings include the microvascular deposition of hyaline fibrin thrombi, thrombocytopenia, microangiopathic hemolytic anemia, fever, renal failure, fluctuating levels of consciousness, and evanescent focal neurologic deficits. The presence of hyaline thrombi in arterioles, capillaries, and venules without any inflammatory changes in the vessel wall is diagnostic. The presence of a severe Coombs-negative hemolytic anemia with schistocytes or fragmented red blood cells in the peripheral blood smear, coupled with thrombocytopenia, and minimal activation of the coagulation system help to confirm the clinical suspicion of TTP. This disorder should be distinguished from vasculitis andSLE, which can predispose patients to TTP. Platelet-associated IgG and complement levels are usually normal in TTP.

The treatment of acuteTTP has focused on the use of exchange transfusion or intensive plasmapheresis coupled with infusion of fresh frozen plasma. Patients with TTP become transiently deficient in a plasma enzyme that depolymerizes ultra-high-molecular-weightvWFreleased from endothelial cells. Therapy may remove abnormal forms of vWF and replenish the deficient enzyme. Overall mortality has been markedly reduced, and the majority of patients with TTP recover from this formerly fatal disorder. Most patients surviving the acute illness recover completely, with no residual renal or neurologic disease. Occasional patients with a chronic, relapsing form of TTP require maintenance plasmapheresis and plasma infusion, and a few patients are controlled only with glucocorticoids.

HEMOLYTIC-UREMIC SYNDROME

HUSis a disease of infancy and early childhood that closely resemblesTTP. Patients present with fever, thrombocytopenia, microangiopathic hemolytic anemia, hypertension, and varying degrees of acute renal failure. In many cases, onset is preceded by a minor febrile or viral illness, and an infectious or immune complex-mediated cause has been proposed. Epidemics related to infection with a specific strain of *Escherichia coli* (O157:H7) have been documented. As in TTP, disseminated intravascular coagulation is not found. In contrast to TTP, the disorder remains localized to the kidney, where hyaline thrombi are seen in the afferent arterioles and glomerular capillaries. Such thrombi are not present in other vessels, and neurologic symptoms, other than those associated with uremia, are uncommon. No therapy is proven effective; however, with dialysis for acute renal failure, the initial mortality is only 5%. Between 10 and 50% of patients have some chronic renal impairment.

HENOCH-SCHONLEIN PURPURA

Henoch-Schonlein, or anaphylactoid, purpura is a distinct, self-limited type of vasculitis that occurs in children and young adults. Patients have an acute inflammatory reaction in capillaries, mesangial tissues, and small arterioles that leads to increased vascular permeability, exudation, and hemorrhage. Vessel lesions contain IgA and complement components. The syndrome may be preceded by an upper respiratory infection or streptococcal pharyngitis or be associated with food or drug allergies. Patients develop a purpuric or urticarial rash on the extensor surfaces of the arms and legs and on the buttocks as well as polyarthralgias or arthritis, colicky abdominal pain, and hematuria from focal glomerulonephritis. Despite the hemorrhagic features, all coagulation tests are normal. A small number of patients may develop fatal acute renal failure, and 5 to 10% develop chronic nephritis. Glucocorticoids provide symptomatic relief of the joint and abdominal pains but do not alter the course of the illness.

METABOLIC AND INFLAMMATORY DISORDERS

Acute febrile illnesses may cause capillary fragility and skin bleeding. Immune complexes containing viral antigens or the viruses themselves may damage endothelial cells. In addition, certain pathogens such as the rickettsiae that cause Rocky Mountain spotted fever replicate in endothelial cells and damage them. Thrombocytopenia is also a frequent finding in acute infectious disorders and may contribute to skin bleeding. In addition, whenever the platelet count is<10,000/uL, gaps develop between endothelial cells, which allow the diapedesis of red cells into the dermis, forming petechiae. Drugs such as the sulfonamides, penicillin, and allopurinol may cause vascular inflammation, resulting in maculopapular or urticarial rashes. Some of these mechanisms are additive, and drug reactions in thrombocytopenic individuals cause an intensely hemorrhagic rash.

Occasionally, patients with diffuse polyclonal hyperglobulinemia will develop purpuric lesions on the lower limbs -- a benign condition referred to as *hyperglobulinemic purpura.* Vascular purpura may occur in patients with various monoclonal gammopathies, including Waldenstrom's macroglobulinemia, multiple myeloma, and cryoglobulinemia. These proteins markedly increase serum viscosity and may impair blood flow through capillaries and lead to retinal hemorrhage, central nervous system dysfunction, and skin necrosis. In addition, the globulins may impair platelet aggregation and adhesion and interfere with fibrin polymerization. Patients with mixed cryoglobulinemia develop a more extensive maculopapular lesion due to immune complex-mediated damage to the vessel wall. The mixed cryoglobulinemia (usually IgG and anti-IgG) may be associated with arthralgias, diffuse weakness, and unexplained nephritis. Plasmapheresis will temporarily lower the level of globulins, remove immune complexes, and improve symptoms in these patients. However, long-term management must include control of the underlying disease that produces the abnormal globulins or immune complexes.

Patients with *scurvy* (vitamin C deficiency) develop painful episodes of perifollicular skin bleeding as well as bleeding into muscles and, occasionally, into the gastrointestinal and genitourinary tracts. The diagnosis is confirmed by the presence of hyperkeratosis of skin, gum swelling, and low levels of the vitamin in leukocytes. Vitamin C is needed to

synthesize hydroxyproline, an essential constituent of collagen. Thus, collagen synthesis is impaired by scurvy. Patients with *Cushing's syndrome*, an excess production of glucocorticoids, or patients on large doses of glucocorticoids develop generalized protein wasting and may show skin bleeding or easy bruising due to atrophy of the supporting connective tissue around blood vessels. Aging causes a similar atrophy of perivascular connective tissue on the extensor surfaces of the hands and arms, leading to *senile purpura* -- dark purple, irregularly shaped hemorrhagic areas due to abnormal skin mobility that tears small blood vessels.

Patients with inherited disorders of the connective tissue matrix such as *Marfan's syndrome*, *Ehlers-Danlos syndrome*, and *pseudoxanthoma elasticum* also have easy bruising. In addition to having fragile skin vessels and easy bruising, patients with Ehlers-Danlos syndrome may develop aneurysms in intraabdominal vessels and apoplectic rupture and hemorrhage due to defects in the vascular collagen network. Primary vascular abnormalities can also lead to bleeding. Patients with *Osler-Weber-Rendu disease* (hereditary hemorrhagic telangiectasia), an inherited autosomal dominant disorder, have frequent episodes of nasal and gastrointestinal bleeding from abnormal telangiectatic capillaries. They may develop pulmonary arteriovenous fistulas. Two genetic defects have been identified in these patients both involving proteins that bind to transforming growth factor b (TGF-b); HHT-1 has mutations in endoglin, and HHT-2 has mutations in ALK-1. Patients with *angiodysplasia* of the colon have increased incidence of gastrointestinal bleeding. In the *Kasabach-Merritt syndrome*, patients may have very extensive and progressively enlarging vascular malformation that may involve large portions of their extremities. Bleeding is secondary to disseminated intravascular coagulation triggered by stagnant blood flow through the tortuous vessels.

(Bibliography omitted in Palm version)

Back to Table of Contents

117. DISORDERS OF COAGULATION AND THROMBOSIS - *Robert I. Handin*

Patients with congenital plasma coagulation defects characteristically bleed into muscles, joints, and body cavities hours or days after an injury. Most of the *inherited* plasma coagulation disorders are due to defects in single coagulation proteins, with the two X-linked disorders, factors VIII and IX deficiency, accounting for the majority. These patients may have severe bleeding and chronic disability and require specialized medical therapy. With rare exceptions, the known disorders prolong either the prothrombin time (PT), partial thromboplastin time (PTT), or both. If they are abnormal, quantitative assays of specific coagulation proteins are then carried out using the PT or PTT tests with plasma from congenitally deficient individuals as substrate. The corrective effect of varying concentrations of patient plasma is measured and expressed as a percentage of a normal pooled plasma standard. The interval range for most coagulation factors is from 50 to 150% of this average value, and the minimal level of most individual factors needed for adequate hemostasis is 25%.

Acquired coagulation disorders are both more frequent and more complex, arising from deficiencies of multiple coagulation proteins and simultaneously affecting both primary and secondary hemostasis. The most common acquired hemorrhagic disorders are (1) disseminated intravascular coagulation (DIC), (2) the hemorrhagic diathesis of liver disease, and (3) vitamin K deficiency and complications of anticoagulant therapy.

Although congenital and acquired bleeding disorders are relatively rare, venous and arterial thrombosis and embolism are common medical disorders that have been recognized for>100 years. Although risk factors such as atherosclerotic vascular disease, congestive heart failure, malignancy, and immobility predispose patients to thrombosis, specific coagulation defects have not yet been identified in most patients with thromboembolism. Several inherited coagulation abnormalities induce a hypercoagulable or prethrombotic state and predispose patients to thrombosis. These disorders affect young people, cause recurrent episodes of thromboembolism, and may involve multiple members of a single family. An understanding of the biochemical basis of thromboembolism is also important because anticoagulant and antithrombotic regimens are based on the premise that modifying critical coagulation reactions will reduce the incidence of thrombosis.**For further discussion of the physiology of normal hemostasis and the cardinal manifestations of the hemorrhagic and thrombotic disorders, see Chap. 62.*

FACTOR VIII DEFICIENCY -- HEMOPHILIA A

Pathogenesis and Clinical Manifestations The antihemophilic factor (AHF), or factor VIII coagulant protein, is a large (265-kDa), single-chain protein that regulates the activation of factor X by proteases generated in the intrinsic coagulation pathway (Figs. 62-5 and62-6). It is synthesized in liver and circulates complexed to the von Willebrand factor (vWF) protein. Factor VIII molecule is present in low concentration (10 ug/L) and is susceptible to proteolysis. The gene for factor VIII is on the X chromosome, and carrier detection and prenatal diagnosis are well established.

One in 10,000 males is born with deficiency or dysfunction of the factor VIII molecule. The resulting disorder, *hemophilia A*, is characterized by bleeding into soft tissues,

muscles, and weight-bearing joints. Symptomatic patients usually have factor VIII levels<5%, with a close correlation between the clinical severity of hemophilia and plasmaAHFlevel. Patients with <1%factor VIII activity have *severe* disease; they bleed frequently even without discernible trauma. Patients with levels of 1 to 5% have *moderate* disease with less frequent bleeding episodes. Those with levels>5% have *mild* disease with infrequent bleeding that is usually secondary to trauma. Occasional patients with factor VIII levels as high as 25% are discovered when they bleed after major trauma or surgery. The majority of patients with hemophilia A have factor VIII levels below<5%.

Hemophilic bleeding occurs hours or days after injury, can involve any organ, and, if untreated, may continue for days or weeks. This can result in large collections of partially clotted blood putting pressure on adjacent normal tissues and can cause necrosis of muscle (compartment syndromes), venous congestion (pseudophlebitis), or ischemic damage to nerves. Patients with hemophilia often develop femoral neuropathy due to pressure from an unsuspected retroperitoneal hematoma. They can also develop large calcified masses of blood and inflammatory tissue that are mistaken for cancers (pseudotumor syndrome).

Patients with severe hemophilia are usually diagnosed shortly after birth because of an extensive cephalhematoma or profuse bleeding at circumcision. However, young children with moderate disease may not bleed until they begin to walk or crawl, and individuals with mild hemophilia may not be diagnosed until they are adolescents or young adults. Typically, a hemophilia patient presents with pain followed by swelling in a weight-bearing joint, such as the hip, knee, or ankle. The presence of blood in the joint (*hemarthrosis*) causes synovial inflammation, and repetitive bleeding erodes articular cartilage and causes osteoarthritis, articular fibrosis, joint ankylosis, and eventually muscle atrophy. Bleeding may occur into any joint, but after a joint has been damaged, it may become a site for subsequent bleeding episodes.

Hematuria, without any genitourinary pathology, is also common. It is usually self-limited and may not require specific therapy. The most feared complications of hemophilia are oropharyngeal and central nervous system bleeding. Patients with oropharyngeal bleeding may require emergency intubation to maintain an adequate airway. Central nervous system bleeding can occur without antecedent trauma or without evidence of a specific lesion.

Patients suspected of having hemophilia should have a platelet count, bleeding time,PT, andPTT. Typically, the patient will have a prolonged PTT with all other tests normal. Because of the clinical similarity of factor VIII deficiency and factor IX deficiency, any male with an appropriate bleeding history and a prolonged PTT should have specific assays for factor VIII and factor IX.

TREATMENT

Tenets regarding the treatment of bleeding in hemophilia patients include the following: (1) Symptoms often precede objective evidence of bleeding. (2) Signs of bleeding may not appear until several days after well-documented trauma. The patients can generally be relied upon to identify early symptoms, usually pain. Early treatment is more

effective, less costly, and can be lifesaving. (3) Avoid the use of aspirin or aspirin-containing drugs, which impair platelet function and may cause severe hemorrhage. COX-2 inhibitors can be used, as they do not impair platelet function.

Plasma products enriched in factor VIII have revolutionized the care of hemophilia patients, reduced the degree of orthopedic deformity, and permitted virtually any form of elective and emergency surgery. The widespread use of factor VIII concentrates has also produced serious complications, including viral hepatitis, chronic liver disease, and AIDS. *Cryoprecipitate*, which contains about half the factor VIII activity of fresh-frozen plasma in one-tenth the original volume, is simple to prepare and is produced in hospital or regional blood banks.

Three developments have increased the safety of factor VIII therapy and have changed medical practice. First, heating of lyophilized factor VIII concentrates under carefully controlled conditions can inactivate HIV without destroying factor VIII activity. Second, highly purified factor VIII can be produced by adsorbing and eluting factor VIII from monoclonal antibody columns. Third, recombinant factor VIII is now available. Patients with hemophilia should receive either monoclonal purified or recombinant factor VIII to minimize viral infections and exposure to irrelevant proteins.

Each unit of factor VIII infused, defined as the amount present in 1 mL normal plasma, will raise the plasma level of the recipient by 2%/kg of body weight. Factor VIII has a half-life of 8 to 12 h, making it necessary to infuse it continuously or at least twice daily to sustain a chosen factor VIII level. In patients with mild hemophilia, an alternative treatment is desmopressin (DDAVP), which transiently increases the factor VIII level. DDAVP will increase the factor level two- to threefold. Although generally safe, it occasionally causes hyponatremia or may precipitate thrombosis in elderly patients.

An uncomplicated episode of soft tissue bleeding or an early hemarthrosis can be treated with one infusion of sufficient factor VIII concentrate to raise the factor VIII level to 15 or 20%. A more extensive hemarthrosis or retroperitoneal bleeding requires twice-daily or continuous infusions in order to keep the factor VIII level at 25 to 50% for at least 72 h. Life-threatening bleeding into the central nervous system or major surgery may require therapy for 2 weeks with levels kept at a minimum of 50% normal. Patients also need skilled orthopedic care, with immobilization of inflamed joints to promote healing and to prevent contractures, and physical therapy to strengthen muscles and maintain joint mobility.

Before surgery, every hemophilia patient should be screened for the presence of an inhibitor to factor VIII. Patients with hemophilia who do not have an inhibitor should receive factor VIII infusions just before surgery and will require daily monitoring so that the factor VIII level is maintained>50% for 10 to 14 days after surgery. When patients undergo joint replacement or other major orthopedic surgery, therapy should be continued for 3 weeks to permit wound healing and the institution of physical therapy.

Hemophilia patients also require treatment before dental procedures. Filling of a carious tooth can be managed by a single infusion of factor VIII concentrate coupled with the administration of 4 to 6 g of e-aminocaproic acid (EACA) four times daily for 3 to 4 days after the dental procedure. EACA is a potent antifibrinolytic agent that inhibits

plasminogen activators present in oral secretions and stabilizes clot formation in oral tissue. Alternatives include tranexamic acid, a longer-acting antifibrinolytic. EACA is also effective when used as a mouthwash. For major oral and periodontal surgery and extractions of permanent teeth, patients should probably be hospitalized briefly and also treated with factor VIII concentrates. Therapy should begin just before surgery and continue for at least 2 to 3 days.

Many centers have organized home-care programs so that patients can administer their own factor VIII infusions with the onset of symptoms. Occasional patients with very frequent bleeding receive regularly scheduled infusions. Despite the expense and inconvenience of "prophylactic" infusions, their use in early childhood has reduced or eliminated hemarthroses. Concern about transmission of AIDS has made some patients reluctant to treat themselves, despite the fact that current blood products carry a very low or no risk of transmitting HIV.

The prospects for correcting factor VIII deficiency by gene therapy are promising; some success has been achieved in dogs. Clinical studies in humans are underway.

Complications Most hemophilia patients have had multiple episodes of hepatitis, and a majority have elevated hepatocellular enzyme levels and abnormalities on liver biopsy. Ten to 20% of patients also have hepatosplenomegaly, and a small number develop chronic active or persistent hepatitis or cirrhosis. A few patients with hemophilia and end-stage liver disease have received liver transplants with cure of both diseases. Along with homosexuals and intravenous drug abusers, hemophilia patients are at high risk for AIDS because they frequently receive blood products; they can also present with the full range of AIDS-related syndromes, including diffuse lymphadenopathy and immune thrombocytopenia. Although up to 50% of multiply transfused hemophiliacs are HIV-positive and many have clinical AIDS, the advances in factor VIII concentrate production should prevent future HIV infection.

Despite frequent bleeding, severe iron-deficiency anemia is uncommon because most of the bleeding is internal and iron is effectively recycled. Mild iron deficiency from chronic epistaxis or gastrointestinal bleeding occurs in some patients. In addition, some patients have developed a mild Coombs-positive hemolytic anemia due to small amounts of anti-A and anti-B antibody that are present in intermediate purity factor VIII concentrates.

Following multiple transfusions, 10 to 20% of patients with severe hemophilia develop inhibitors to factor VIII. Inhibitors are usually IgG antibodies that rapidly neutralize factor VIII activity. Two types of inhibitors are found with different biologic characteristics and different clinical presentations. Patients with type I inhibitors have a typical anamnestic response and raise their antibody titer following exposure to factor VIII. Patients with a type II inhibitor have a low antibody titer that is not stimulated by factor VIII infusion. Patients with the type I inhibitor should not receive factor VIII. Control of bleeding may require the infusion of either porcine factor VIII concentrates, which may not be affected by inhibitors, or prothrombin complex concentrates, which contain trace quantities of activated coagulation factors and can bypass the block in coagulation produced by the inhibitor. Patients with low-titer type II antibodies may respond to higher doses of factor VIII.

Protocols to induce tolerance to human factor VIII use massive doses of the factor coupled with immunosuppression. Tolerance induction is expensive and not always effective; it should be reserved for severely affected patients.

Genetic Counseling and Carrier Detection It is possible to trace the defective allele in some families by examining the inheritance of restriction fragment length polymorphisms (RFLP) linked to the factor VIII gene. In addition, in families in which a specific mutation has been defined in the factor VIII gene, it can be readily detected by gene amplification and allele-specific oligonucleotide hybridization. For example, 45% of patients with severe hemophilia A have a chromosomal inversion arising from homologous recombination between sequences in intron 22 and an upstream gene. The inversion is readily detected by polymerase chain reaction (PCR) or Southern blotting. Precise diagnosis is possible early in pregnancy from either chorionic villus biopsy or amniocentesis.

Female carriers of hemophilia, who are heterozygotes, usually produce sufficient factor VIII from the factor VIII allele on their normal X chromosome for normal hemostasis. However, occasional hemophilia carriers will have factor VIII levels far below 50% due to random inactivation of normal X chromosomes in tissue producing factor VIII. These symptomatic carriers may bleed with major surgery or bleed occasionally with menses. Rarely, true female hemophiliacs arise from consanguinity within families with hemophilia or from concomitant Turner's syndrome or XO mosaicism in a carrier female.

FACTOR IX DEFICIENCY -- HEMOPHILIA B

Factor IX is a single-chain, 55-kDa proenzyme that is converted to an active protease (IXa) by factor XIa or by the tissue factor-VIIa complex. Factor IXa then activates factor X in conjunction with activated factor VIII. Factor IX is one of six proteins synthesized in the liver that require vitamin K for biologic activity. Vitamin K is a cofactor for a unique posttranslational modification that inserts a second carboxyl group onto certain glutamic acid residues on factor IX (Chap. 62). This modification permits calcium binding and adsorption onto phospholipid surfaces. Factor IX gene is on the X chromosome.

Factor IX deficiency or dysfunction (hemophilia B, Christmas disease) occurs in 1 in 100,000 male births. Accurate laboratory diagnosis is critical, since it is indistinguishable clinically from factor VIII deficiency (hemophilia A) but requires different treatment. Either fresh-frozen plasma or a plasma fraction enriched in the prothrombin complex proteins is used. Monoclonally purified or recombinant factor IX preparations are now available. In addition to the expected complications of hepatitis, chronic liver disease, and AIDS, the therapy of factor IX deficiency has a special hazard. Trace quantities of activated coagulation factors in prothrombin complex concentrates may activate the coagulation system and cause thrombosis and embolism. This is particularly common in immobilized surgical patients and patients with liver disease. As a result, some centers have returned to fresh-frozen plasma for factor IX-deficient surgical patients, while others have recommended the addition of small doses of heparin to the concentrate to activate antithrombin III during the infusion and reduce hypercoagulability. The recombinant or monoclonally purified products are less likely to be thrombogenic.

FACTOR XI DEFICIENCY

Factor XI is a 160-kDa dimeric protein activated to an active protease (XIa) by factor XIIa, in conjunction with high-molecular-weight kininogen and kallikrein (Figs. 62-5 and62-6). Factor XI deficiency is inherited as an autosomal recessive trait and is especially common in Ashkenazi Jews. In contrast to deficiency in factors VIII and IX, the correlation between factor level and propensity to bleed is not as precise, spontaneous bleeding is less, and hemarthroses are rare. Many patients with factor XI deficiency present with posttraumatic bleeding or with bleeding in the perioperative period, and occasional factor XI-deficient women have menorrhagia. Daily infusions of fresh-frozen plasma are sufficient, since the half-life of factor XI is approximately 24 h. The majority of defective factor XI alleles were accounted for by a limited number of mutations.

OTHER FACTOR DEFICIENCIES

Deficiencies in factors V, VII, X, and prothrombin (factor II) are exceedingly rare autosomal recessive disorders. Spontaneous or posttraumatic musculoskeletal bleeding or menorrhagia can occur with these deficiencies, but hemarthroses are uncommon. Fresh-frozen plasma is the appropriate therapy, although prothrombin concentrates may be employed for patients with severe prothrombin deficiency or decreases in factors VII and X as long as the risks of hepatitis and thrombosis are recognized.

Defects in the contact activation pathway involving Hageman factor (factor XII), high-molecular-weight kininogen, and prekallikrein cause laboratory abnormalities but no clinical bleeding. Despite dramatic prolongation of thePTT, often to greater than 100 s, deficient individuals have normal hemostasis and can undergo major surgery without plasma replacement therapy. Direct activation of factor IX by the tissue factor-VIIa complex may bypass this defective step in coagulation (Fig. 62-7). Recognition of these disorders is important because such patients should neither be treated inappropriately with plasma nor denied indicated surgery on the basis of these laboratory abnormalities.

AFIBRINOGENEMIA AND DYSFIBRINOGENEMIA

Fibrinogen is a 340-kDa dimeric molecule made up of two sets of three covalently linked polypeptide chains. Thrombin sequentially cleaves fibrinopeptides A and B from the Aa and Bb chains of fibrinogen to produce fibrin monomer, which then polymerizes to form a fibrin clot. Although fibrinogen is needed for platelet aggregation and fibrin formation, severe fibrinogen deficiency does not usually cause serious bleeding except after surgery. Patients with afibrinogenemia, who have no detectable fibrinogen in plasma or platelets, may have infrequent, mild bleeding episodes. Preliminary genetic analyses do not show any gross deletion or structural changes in the genes encoding the a,b, and g chains of fibrinogen despite the total absence of plasma fibrinogen.

Fibrinogen is an abundant plasma protein (2.5 g/L). Mutations have been identified that alter the release of fibrinopeptides from the Aa and Bb chains of fibrinogen, the rate of polymerization of fibrin monomers, and the sites for fibrin cross-linking. These dysfibrinogenemias are almost always inherited as autosomal dominant traits, so patients have nearly equal concentrations of normal and mutant fibrinogen in their

plasma. Patients with dysfibrinogenemia have a slightly prolongedPT andPTT, a prolonged thrombin time, and a disparity in levels of fibrinogen measured with functional and immunologic assays. Despite these abnormalities, most patients have no symptoms or only moderate bleeding. A few dysfibrinogenemias induce a hypercoagulable state and increase the risk of thrombosis, and others have been associated with an increased incidence of abortion (Chap. 118). Some patients with liver disease, hepatomas, AIDS, and lymphoproliferative disorders develop an acquired form of dysfibrinogenemia.

FACTOR XIII DEFICIENCY AND DEFECTIVE FIBRIN CROSS-LINKING

Factor XIII is a transglutaminase that stabilizes fibrin clots by forminge-amino-g-glutamyl cross-links between adjacent a andg chains of fibrin. Factor XIII deficiency is an extremely rare inherited syndrome. Patients usually bleed in the neonatal period from their umbilical stump or circumcision. In addition to hemorrhage, these patients may have poor wound healing, a high incidence of infertility among males and abortion among affected females, and a high incidence of intracerebral hemorrhage. These observations suggest that the enzyme may be important in other physiologic and pathologic processes beyond hemostasis, including placental implantation, spermatogenesis, and wound healing. Several drugs, including isoniazid, may bind to cross-linking sites on fibrinogen and mimic factor XIII deficiency by blocking enzyme activity. Normal hemostasis requires only 1% of normal enzyme activity, which can be achieved with a single infusion of fresh-frozen plasma or a purified factor XIII-rich product derived from human placenta called Fibrogammin. Factor XIII has a 14-day half-life.

VITAMIN K DEFICIENCY

Vitamin K is a fat-soluble vitamin that plays a critical role in hemostasis. Dietary vitamin K is absorbed in the small intestine and stored in the liver. The vitamin is also synthesized by endogenous bacterial flora in the small intestine and colon; however, the quantity of endogenous vitamin K absorbed from the large intestine is debated. Following absorption and transport, vitamin K is converted to an active epoxide in liver microsomes and serves as a cofactor in the enzymatic carboxylation of glutamic acid residues on prothrombin complex proteins (Fig. 117-1).

The three major causes of vitamin K deficiency are inadequate dietary intake, intestinal malabsorption, and loss of storage sites due to hepatocellular disease. Neonatal vitamin K deficiency, which causes hemorrhagic disease of the newborn, has disappeared from western countries with the routine administration of vitamin K to all newborn infants. Although a 30-day supply of vitamin K is stored in the normal liver, acutely ill patients can become deficient within 7 to 10 days. Acute vitamin K deficiency is particularly common in patients recovering from biliary tract surgery who have no dietary intake of vitamin K, have T-tube drainage of bile, and are on broad-spectrum antibiotics. Vitamin K deficiency is also seen in chronic liver disease, particularly primary biliary cirrhosis, and in some malabsorption states (Chaps. 286 and298). The cephalosporins inhibit the reduction and recycling of vitamin K, much like coumarin.

With vitamin K deficiency, plasma levels of all the prothrombin complex proteins (factors II, VII, IX, X; proteins C and S) decrease. Factor VII and protein C, which have the

shortest half-lives, decrease first. Because of the rapid fall in factor VII, patients with mild vitamin K deficiency may have a prolongedPT and a normalPTT. Later, as the levels of the other factors fall, the PTT will also become prolonged. Parenteral administration of 10 mg vitamin K rapidly restores vitamin K levels in the liver and permits normal production of prothrombin complex proteins within 8 to 10 h. Severe hemorrhage can be treated with fresh-frozen plasma, which immediately corrects the hemostatic defect. If the cause of vitamin K deficiency cannot be eliminated, patients may need monthly injections. Purified prothrombin complex concentrates should be avoided because they contain trace quantities of activated forms of the prothrombin complex proteins and can cause thrombosis in patients with liver disease. They also carry an increased risk of hepatitis.

DISSEMINATED INTRAVASCULAR COAGULATION

DICcan be either an explosive and life-threatening bleeding disorder or a relatively mild or subclinical disorder. Although a long list of diseases can be complicated by DIC, it is most frequently associated with obstetric catastrophes, metastatic malignancy, massive trauma, and bacterial sepsis (Table 117-1). In each case, a tentative triggering mechanism has been identified. For example, tumors and traumatized or necrotic tissue release tissue factor into the circulation, while endotoxin from gram-negative bacteria activates several steps in the coagulation cascade. In addition to a direct effect on the activation of Hageman factor (factor XII), endotoxin induces the expression of tissue factor on the surface of monocytes and endothelial cells. These activated cell surfaces then accelerate coagulation reactions. These potent thrombogenic stimuli cause the deposition of small thrombi and emboli throughout the microvasculature. This early thrombotic phase of DIC is then followed by a phase of procoagulant consumption and secondary fibrinolysis. Continued fibrin formation and fibrinolysis lead to hemorrhage from the coagulation factor and platelet depletion and the antihemostatic effects of fibrin degradation products (Fig. 117-2).

The clinical presentation varies with the stage and severity of the syndrome. Most patients have extensive skin and mucous membrane bleeding and hemorrhage from surgical incisions or venipuncture or catheter sites. Less often, patients present with peripheral acrocyanosis, thrombosis, and pregangrenous changes in digits, genitalia, and nose -- areas where blood flow is markedly reduced by vasospasm or microthrombi. Some patients, particularly those with chronicDICsecondary to malignancy, have laboratory abnormalities without any evidence of thrombosis or hemorrhage.

The laboratory manifestations include thrombocytopenia and the presence of schistocytes or fragmented red blood cells that arise from cell trapping and damage within fibrin thrombi; prolongedPT andPTT and thrombin time and a reduced fibrinogen level from depletion of coagulation proteins; and elevated fibrin degradation products (FDP) from intense secondary fibrinolysis. The D dimer immunoassay, which specifically measures cross-linked fibrin derivatives, is a more specific FDP assay. The abnormality inDIC that predicts bleeding is the plasma fibrinogen level; low fibrinogen levels are associated with more bleeding.

TREATMENT

DIC, although sometimes indolent, can cause life-threatening hemorrhage and may require emergency treatment. This should include (1) an attempt to correct any reversible cause of DIC; (2) measures to control the major symptom, either bleeding or thrombosis; and (3) a prophylactic regimen to prevent recurrence in cases of chronic DIC. Treatment will vary with the clinical presentation. In patients with an obstetric complication such as abruptio placentae or acute bacterial sepsis, the underlying disorder is easy to correct, and prompt delivery of the fetus and placenta or treatment with appropriate antibiotics will reverse the DIC syndrome. In patients with metastatic tumor causing DIC, control of the primary disease may not be possible, and long-term prophylaxis may be necessary.

Patients with bleeding as a major symptom should receive fresh-frozen plasma to replace depleted clotting factors and platelet concentrates to correct thrombocytopenia. Those with acrocyanosis and incipient gangrene or other thrombotic problems need immediate anticoagulation with intravenous heparin. The use of heparin in the treatment of bleeding is still controversial. Although it is a logical way to reduce thrombin generation and prevent further consumption of clotting proteins, it should be reserved for patients with thrombosis or who continue to bleed despite vigorous treatment with plasma and platelets.

Patients who initially have mildDIC and may not be symptomatic may begin to bleed following surgery or chemotherapy. For example, mild DIC, without clinical bleeding, has been documented during saline- or prostaglandin-induced midtrimester abortions. Prophylactic treatment of patients with heparin may prevent progression of a mild DIC syndrome and has been used in the treatment of patients with acute promyelocytic leukemia and in some patients with a retained dead fetus who require surgical extraction. However, most patients with low-grade DIC can be managed with plasma and platelet replacement and do not require heparin. Chronic DIC does not respond to oral warfarin anticoagulants, but it can be controlled with long-term heparin infusion. Occasional patients with indolent tumors and severe DIC have been maintained on heparin administered by intermittent subcutaneous injection or continuous infusion with portable pumps.

Despite our detailed understanding of the pathophysiology ofDIC and a vigorous approach to therapy, treatment does not change the natural history of the underlying disorder. Therapy will only stabilize the patient, prevent exsanguination or massive thrombosis, and permit institution of definitive therapy.

COAGULATION DISORDERS IN LIVER DISEASE

Liver dysfunction is frequently accompanied by a hemostatic defect. The major causes of hemorrhage in patients with liver disease are shown in Table 117-2. Bleeding is usually due to an anatomic lesion that is exacerbated by a hemostatic defect. Most patients bleed from complications of portal hypertension, esophageal varices, or gastritis and peptic ulcer disease. Portal hypertension also causes splenomegaly, with splenic sequestration of platelets and thrombocytopenia, which contributes to the hemostatic defect (Chap. 298).

Patients with hepatocellular liver disease cannot store vitamin K optimally and may have

some degree of vitamin K deficiency. Cholestasis, a frequent feature of liver disease, impairs vitamin K absorption and further decreases liver vitamin K stores. Abnormalities in the g-carboxylation of prothrombin complex proteins independent of vitamin K and the production of abnormal proteins have also been described. Patients may also have decreased production of other coagulation proteins, including fibrinogen and factor V. The liver also produces inhibitors of coagulation such as antithrombin III and proteins C and S and is the clearance site for activated coagulation factors and fibrinolytic enzymes. Thus patients with liver disease are also "hypercoagulable" and predisposed to developingDIC or systemic fibrinolysis. Coagulation defects in advanced liver failure are often difficult to distinguish from those of DIC.

Each patient with hemorrhage and liver disease should have aPT,PTT, platelet count, and fibrinogen determination, although it is not always possible to determine the major hemostatic abnormality from a single set of laboratory values. It is helpful to have previous laboratory data available for patients with chronic liver disease who develop an acute complication. The degree of prolongation of the PT predicts the risk of bleeding. Most patients present with moderate prolongation of the PT and PTT, mild thrombocytopenia, and a normal fibrinogen level. However, they may also present with a more complex defect combining defective synthesis, abnormal clearance, and active consumption of coagulation proteins. Since vitamin K deficiency is so common, a single parenteral dose of vitamin K is given after initial laboratory studies have been obtained, even though this may only partially correct the laboratory abnormalities. The presence of severe thrombocytopenia or a low fibrinogen level suggests the additional complication ofDIC and may require further studies and therapy.

The safest replacement therapy for a patient with liver disease is fresh-frozen plasma, since it supplies all known coagulation factors. However, even this form of therapy has drawbacks, since large quantities of plasma may precipitate hepatic encephalopathy and cause fluid and sodium overload. Prothrombin complex concentrates should be avoided because they replace only the vitamin K-dependent factors, may be contaminated with hepatitis and AIDS virus, and contain trace quantities of activated coagulation proteins. Similarly, fibrinogen concentrates (or cryoprecipitate), rich in factor VIII and fibrinogen, should not be used without additional fresh-frozen plasma. Anticoagulation with heparin has been advocated to controlDIC, but this is particularly hazardous and not recommended in cirrhosis because heparin is metabolized erratically and may lead to severe bleeding.

FIBRINOLYTIC DEFECTS

Bleeding can also occur from defects in the fibrinolytic system. Patients with$\alpha_2$plasmin inhibitor deficiency or plasminogen activator inhibitor (PAI) 1 have rapid fibrinolysis following fibrin deposition after trauma or surgery and may experience recurrent hemorrhage. Similarly, patients with cirrhosis have an impaired clearance of tissue plasminogen activator (tPA) and systemic fibrinolysis that may contribute to their hemorrhagic defect. Rarely, patients with tumors such as metastatic prostatic cancer may develop diffuse bleeding from primary fibrinolysis rather thanDIC. Clues to the diagnosis include a disproportionately low fibrinogen level with a relatively normalPT andPTT and the presence of a normal or nearly normal platelet count. With rare exceptions, patients with primary fibrinolysis should have an elevated titer ofFDP but a

normal D dimer level. However, it is sometimes difficult or impossible to differentiate primary fibrinolysis from the secondary fibrinolysis accompanying DIC. Patients with clearly established primary fibrinolysis should not receive heparin; they require plasma therapy and, occasionally, fibrinolytic inhibitors such asEACA. However, EACA should not be given to patients suspected of having DIC unless they are also receiving heparin, since EACA can cause massive, often fatal, thrombosis in a patient with DIC.

CIRCULATING ANTICOAGULANTS

Circulating anticoagulants, or inhibitors, are usually IgG antibodies that interfere with coagulation reactions. Specific inhibitors inactivate individual coagulation proteins and may cause severe hemorrhage. They arise in 15 to 20%of patients with factor VIII or factor IX deficiency who have received plasma infusions. *Specific* inhibitors also occur in previously normal individuals. Although the most common target protein is factor VIII, inhibitors with specificity for each of the coagulation proteins occur. In addition to hemophiliacs, anti-factor VIII antibodies are seen in postpartum females, in patients on various drugs, as part of the spectrum of autoantibodies in systemic lupus erythematosus (SLE) patients, and in normal elderly individuals. Circulating anticoagulants also occur in patients with AIDS.

Nonspecific (lupus-like) inhibitors prolong coagulation tests by binding to phospholipids. They are assayed by their anticoagulant effect [lupus anticoagulant (LA) activity] or their ability to bind to the complex phospholipid cardiolipin [anticardiolipin antibody (ACLA) activity]. While most often encountered in patients withSLE, these nonspecific inhibitors may develop in patients with many other disorders and also in otherwise normal individuals.

The critical laboratory feature that identifies the presence of either type of inhibitor is the failure of normal plasma to correct a prolongedPT,PTT, or both. Plasma from patients with a specific inhibitor will progressively inactivate a coagulation protein and thus prolong whichever of these screening tests requires the participation of that clotting factor. This effect persists after dilution. Nonspecific inhibitors immediately prolong the PT and PTT and, at low dilution, block multiple coagulation reactions. However, these effects can be overcome by altering the quantity or type of phospholipid or by diluting the plasma.

Hemorrhage in patients with specific inhibitors may require treatment with massive plasma or concentrate infusion, the use of activated prothrombin complex concentrates to bypass the antibodies against factors VIII or IX, and plasmapheresis or exchange transfusion to lower antibody titer. Chronic immunosuppressive regimens have been particularly useful in otherwise normal individuals with an acquired factor VIII antibody. Many patients lose their antibody and recover within 6 to 12 months, although the acute mortality rate from uncontrollable bleeding may approach 10%.

Patients withLAactivity have normal hemostasis and will not bleed unless they have concomitant thrombocytopenia or prothrombin deficiency. Both thrombocytopenia and hypoprothrombinemia are secondary to autoantibodies that bind either to platelets or the prothrombin molecule. While these antibodies have no effect on function, they accelerate clearance of the coated platelets or the antibody-prothrombin complexes.

The presence ofLAactivity may predispose patients to venous and arterial thromboembolism and may cause midtrimester abortions. However, the risk of thrombosis is difficult to estimate and the appropriate therapy for individual patients difficult to choose. Tests for either LA orACLAactivity are not well standardized, and results vary among and within patients. The best predictor is a consistent prolongation of more than one coagulation test coupled with a high titer of ACLA activity. Second, the risk of thrombosis is increased in patients who haveSLEcompared with those with idiopathic LA or ACLA activity. Prophylactic therapy is not clearly beneficial, and treatments aimed at reducing the titer of antibody are not superior to conventional antithrombotic therapy.

Therapy should be individualized. Patients withSLE and eitherLA orACLAactivity who have had a thrombotic episode are at high risk for a recurrence and should receive long-term anticoagulant therapy. Women who have had more than one midtrimester abortion, especially those with SLE, should have a trial of anticoagulant therapy. Patients with a single thrombotic episode (stroke or pulmonary embolus) and no other risk factor except LA or ACLA activity should be treated. No consensus has been reached about treatment after an initial minor event [deep venous thrombosis (DVT)]. Asymptomatic patients with only laboratory abnormalities should not be treated. Glucocorticoids should be administered only in conjunction with antithrombotic agents and are not of proven efficacy.

INHERITED PRETHROMBOTIC DISORDERS

Coagulation is carefully regulated by a series of inhibitors that limit thrombin generation and fibrin formation and by the fibrinolytic system, which effectively removes fibrin thrombi (Figs. 62-5 and62-6). Inherited defects in the natural coagulation inhibitors (i.e., antithrombin, proteins C and S), abnormalities in the fibrinolytic system, and certain dysfibrinogenemias predispose patients to thrombosis (Table 62-5). A single point mutation in the factor V gene (factor V Leiden), which converts arginine 506 to glutamine and makes the molecule resistant to degradation by activated protein C, may account for 25% of inherited prethrombotic states. Antithrombin, protein C, and protein S defects are all autosomal dominant traits, so heterozygous individuals, who have a 50% reduction in protein concentration or a mixture of mutant and normal molecules, will have an increased risk of thrombosis. The patients have similar clinical presentations with a strong family history of thrombosis, episodes of recurrent venous thromboembolism, and symptoms by their early twenties. Any patient with this distinctive history should be tested for specific abnormalities.

ANTITHROMBIN DEFICIENCY

Antithrombin complexes with activated coagulation proteins and blocks their biologic activity (Fig. 62-5). The rate of this reaction is enhanced by heparin-like molecules within the vessel wall or on endothelial cells. Plasma antithrombin III content is 5 to 15 mg/L (50 to 150%), with values only slightly below normal increasing the risk of thrombosis. For optimal screening, the antithrombin III concentration is measured by immunoassay and the plasma antithrombin and heparin cofactor activity assessed with functional assays. The most common defect (1 in 2000 indidivuals) is mild (heterozygous)

antithrombin deficiency. Dysfunctional antithrombin molecules with mutations affecting either the serine protease or heparin-binding site or activation of inhibitor by heparin have also been described.

Patients with antithrombin deficiency who develop acute thrombosis or embolism can be treated with intravenous heparin, since there is usually sufficient normal antithrombin to act as a heparin cofactor. Following their first episode of thromboembolism, patients should be placed on oral anticoagulants for life to prevent recurrent thrombosis. Family studies should be conducted when an antithrombin-deficient individual is discovered, since up to half the members of a kindred may be affected. Asymptomatic individuals with antithrombin deficiency should receive prophylactic anticoagulation with heparin or plasma infusions to raise their antithrombin level before medical or surgical procedures that may increase their risk of thrombosis. Chronic oral anticoagulation is not recommended until individuals at risk have a thrombotic episode.

DEFICIENCIES OF PROTEINS C AND S

Protein C is a vitamin K-dependent hepatic protein that binds to the endothelial cell surface protein thrombomodulin and is converted to an active protease by thrombin (Fig. 62-5). Activated protein C, in conjunction with protein S, proteolyzes factors Va and VIIIa, which shuts off fibrin formation. Activated protein C may also stimulate fibrinolysis and accelerate clot lysis. Deficiencies of proteins C and S are usually autosomal dominant disorders, and deficiencies in the two proteins cause an identical syndrome of recurrent venous thrombosis and pulmonary embolism. Dysfunctional molecules have also been identified in some patients with thrombosis. Rare patients with homozygous protein C deficiency have fulminant neonatal intravascular coagulation and require prompt diagnosis and treatment.

The correlation between protein C and S levels and the risk of thrombosis is not as precise as for antithrombin III deficiency. In fact, some asymptomatic individuals with protein C "deficiency" have been discovered. In some well-studied protein C-deficient kindreds, asymptomatic individuals may have protein C levels as low as or lower than relatives with recurrent thrombosis. It is possible that an undiscovered cofactor is present in symptomatic patients. Finally, since a fraction of the available protein S is bound to C4b-binding protein and is unavailable for coagulation reactions, both free and total protein S levels or C4b-binding protein levels should be assessed for maximum accuracy.

Heterozygous patients with protein C or S deficiencies who develop acute thrombosis should be heparinized and then placed on oral anticoagulants. There are, however, two potential problems with the use of coumarin anticoagulants in these patients. First, these vitamin K antagonists (Fig. 117-1;Fig. 62-5), which lower the level of the procoagulant factors II, VII, IX, and X, may also reduce the concentration of proteins C and S sufficiently to nullify the desired antithrombotic effect. In addition, patients who are protein C-deficient may develop coumarin-induced skin necrosis; this defect may predispose patients to a rare but serious complication. Patients with homozygous protein C deficiency require periodic plasma infusions rather than oral anticoagulants to prevent recurrent intravascular coagulation and thrombosis.

RESISTANCE TO ACTIVATED PROTEIN C AND THE FACTOR V LEIDEN MUTATION

Some patients with familial or recurrent venous thromboembolism were found not to prolong theirPTT when activated protein C was added to their plasma. These patients were found to have an identical mutation in which arginine 506 in factor V is converted to glutamine. This amino acid substitution abolishes a protein C cleavage site in factor V and thus prolongs the thrombogenic effect of factor V activation. About 3% of the population worldwide is heterozygous for this mutation. The mutation is absent in certain populations, e.g., Asians, African Americans, and Native Americans. It may account for 25% of patients with recurrent deep venous thrombosis or pulmonary embolism.

Heterozygosity at this allele increases an individual's lifetime risk of venous thromboembolism sevenfold. The risk rises steadily with age. A homozygote has a twentyfold increased risk of thrombosis. Heterozygosity coupled with ingestion of oral contraceptives or pregnancy increases the risk at least fifteenfold. Coinheritance of factor V Leiden and another low-penetrance defect such as protein C or S deficiency is also additive. Many previous studies of risk factors predisposing patients to venous thromboembolism are being reevaluated to take into account this common mutation.

PROTHROMBIN GENE MUTATION

A specific point mutation in the prothrombin gene [conversion of G to A at position 20210 (G20210A)] also predisposes to venous thrombosis and embolism. This mutation is in the 3¢-untranslated region of the gene and results in a 30% increase in plasma prothrombin levels, either through more efficient translation or greater stability of the message. Heterozygotes account for ~18% of cases with family histories of venous thrombosis and 6% of patients with first episodes ofDVT.

The inheritance of multiple mutations increases the risk of thrombosis. The relationship between known mutations and the type of thrombosis is shown inTable 117-3. The fraction of patients withDVT with known mutations is shown in Table 117-4.

TREATMENT

Patients who develop venous thromboembolism without a clear predisposing factor, have a strong family history, present under the age 30, or have more than one episode should have assays for antithrombin III, proteins C and S, and factor V Leiden. Patients who present withDVT or pulmonary embolism during pregnancy or while using oral contraceptives have a 30% chance of having factor V Leiden.

Treatment recommendations for patients with the inherited prethrombotic disorders are still evolving. All patients should receive standard initial therapy with heparin, either conventional or low dose (Chap. 118), followed by 3 months of oral warfarin. This regimen should allow for maximal healing and reendothelialization of the thrombosed vessels and minimize recurrence in the damaged vascular beds. It is not clear which patients should go on to receive long-term (perhaps lifelong) anticoagulation, a judgment that depends on assessing the risk/benefit ratio.

Patients with antithrombin III deficiency who become symptomatic have a high likelihood of recurrent events and should be placed on lifelong anticoagulation. Patients with protein C or S deficiency or heterozygous factor V Leiden and prothrombin G20210A patients have a lower likelihood of recurrent disease. Long-term anticoagulation should be reserved until their second or subsequent episode of thromboembolism. Homozygous factor V Leiden patients should be placed on long-term anticoagulation after their initial episode, and all patients should receive replacement therapy or receive heparin prophylaxis during surgery or after trauma; women with these defects should avoid the use of oral contraceptives. The asymptomatic relatives of patients shown to have these disorders should be screened to determine if they have inherited the defective gene. If so, they should receive appropriate prophylaxis but not start anticoagulation until they are symptomatic. In the absence of a congenital defect predisposing a patient to thrombosis, recurring or migratory thrombophlebitis may indicate an underlying malignancy.

DYSFIBRINOGENEMIAS AND FIBRINOLYTIC DEFECTS

Recurrent venous thrombosis and embolism may be due to familial defects in fibrinogen or plasminogen or decreased synthesis or release oftPA. While most dysfibrinogenemias cause bleeding, several variants have excessively rapid release of fibrinopeptides and recurrent thromboembolism. Patients with this disorder and those with an abnormal plasminogen that resists activation by streptokinase and urokinase have been treated successfully with heparin and oral anticoagulants. Defects in tPA content or release have not been completely characterized. One group of patients with recurrent venous thrombosis and embolism failed to increase venous blood fibrinolytic activity when challenged with local ischemia or physical exercise. The other group had impaired fibrinolytic activity in extracts prepared from biopsied veins. Young patients with acute myocardial infarction may have impaired fibrinolysis due to increased plasma levels ofPAI, a serine protease inhibitor that binds to tPA and is derived from endothelial cells.

Many common illnesses are associated with an increased risk of thrombosis (Table 62-5). These patients are said to have a "hypercoagulable" or "prethrombotic" state. This increased risk is seen in patients with chronic congestive heart failure and metastatic cancer and in patients undergoing major surgery. The generation of tissue factor activity in damaged or ischemic tissue or metastatic tumor, coupled with venous stasis and endothelial injury, induces the formation of venous and, more rarely, arterial thrombi. Several hematologic disorders, paroxysmal nocturnal hemoglobinuria, essential thrombocythemia, and polycythemia vera predispose patients to venous and arterial thrombosis through diverse mechanisms related to increased blood viscosity and abnormal blood cells. Diseases that affect the endothelial cell, such as Behcet's syndrome, Kawasaki's disease, and homocystinuria, or the administration of drugs such as the oral contraceptives, which lower antithrombin III levels, or L-asparaginase, which inhibits production of multiple coagulation factors, may also predispose patients to thrombosis. Infusion of granulocyte-macrophage colony stimulating factor (GM-CSF) has been associated with thrombosis. Tamoxifen, an estrogen receptor antagonist, can cause venous thrombosis. The mechanism is unclear.

Plasma homocysteine levels influence the risk of both venous and arterial

thromboembolism. Individuals with the congenital homocystinuria syndrome have, in addition to their Marfanoid habitus, an increased incidence of strokes and coronary artery disease. These patients have well-recognized enzyme defects (Chap. 352), excrete homocysteine in their urine, and have very high plasma levels of the amino acid. Some patients with early-onset cerebral vascular events have mild homocystinuria that can be brought out by a methionine loading test. Epidemiologic studies show a relationship between homocysteine levels that are nearer to the normal range and coronary artery disease. Although this correlation is not yet definitive, the relationship remains intriguing and of potential clinical relevance. Vitamin B_{12}deficiency occurs in about 30% of people over age 70, produces elevated homocysteine levels, and may be a reversible cause of thrombotic disease.

(Bibliography omitted in Palm version)

Back to Table of Contents

118. ANTICOAGULANT, FIBRINOLYTIC, AND ANTIPLATELET THERAPY - *Robert I. Handin*

ANTICOAGULANT AND FIBRINOLYTIC THERAPY

Anticoagulation with heparin, followed by treatment with oral vitamin K antagonists, is the standard treatment for acute venous thrombosis and pulmonary embolism. In addition, chronic oral anticoagulation is used to prevent cerebral arterial embolism from cardiac sources such as ventricular mural thrombi or atrial thrombi or from an atherosclerotic, partially stenosed carotid or vertebral artery. Anticoagulants are also used, less successfully, to treat peripheral or mesenteric arterial thrombosis. These agents retard fibrin deposition on established thrombi and prevent the formation of new thrombi. The induction of a fibrinolytic state by the infusion of plasminogen activators such as recombinant tissue plasminogen activator (rtPA), streptokinase (SK), or urokinase (UK) has become an accepted mode of therapy for some thromboembolic disorders. Fibrinolytic therapy has been proposed for patients with massive pulmonary embolism and systemic hypotension and to restore the patency of acutely occluded peripheral and coronary arteries. Prompt fibrinolytic therapy can reduce both myocardial damage and mortality following acute coronary occlusion (Chap. 243), though mechanical interventions such as angioplasty and stent placement are also effective ways to restore vessel patency. Fibrinolytic therapy may also be effective in acute thrombotic strokes and in venoocclusive disease of the liver.

ACUTE ANTICOAGULATION WITH HEPARIN

Heparin is a naturally occurring mucopolysaccharide polymer with tetrasaccharide sequences that bind to and activate antithrombin III. It can dramatically reduce thrombin generation and fibrin formation in patients with acute venous and arterial thrombosis or embolism (Table 118-1). Heparin is administered to patients with acute thrombosis or embolism by giving an initial loading dose of 5000 to 10,000 units followed by a continuous intravenous infusion at a rate sufficient to keep the activated partial thromboplastin time (APTT) at 1.5 to 2 times the patient's preheparin APTT. This requires infusion of 800 to 1000 U.S.P. units per hour and is continued while patients are begun on oral anticoagulants and achieve appropriate prolongation of the prothrombin time. The usual duration of combined heparin-warfarin therapy is 5 to 7 days. Heparin is then discontinued, and the patient is maintained on warfarin. Alternatives to a continuous infusion include the administration of 5000 U.S.P. units of heparin four times a day either subcutaneously or intravenously. Unfractionated conventional heparin preparations are heterogeneous, with only 20% of the product biologically active. In addition, active heparin fractions may vary considerably in molecular weight. Biologically active, low-molecular-weight heparin (LMWH) preparations, while more expensive than unfractionated heparin, have several advantages: (1) they can be administered subcutaneously once or twice daily, (2) their pharmacokinetics are so predictable that APTT monitoring is not necessary, and (3) they are less immunogenic and less likely to cause thrombocytopenia. Many patients with deep venous thrombosis, a frequent cause for hospitalization, can be given LMWH as outpatients. Given their other advantages and equivalent efficacy, LMWH preparations have largely replaced unfractionated heparin (Table 118-1).

Patients with recurrent thromboembolism refractory to oral anticoagulants, pregnant women with thromboembolism, and patients with chronic disseminated intravascular coagulation (DIC) may be treated with daily injections ofLMWHpreparations such as enoxoparin or dalteparin. These agents are also effective in prevention of venous thrombosis in high-risk surgical and medical patients, including those with congestive heart failure, myocardial infarction, or cardiomyopathy.

The major complication of unfractionated heparin therapy is bleeding -- especially from surgical sites and into the retroperitoneum. Aspirin or aspirin-containing drugs impair platelet function. Thus, intramuscular injections in patients on both heparin and an antiplatelet drug may cause significant bleeding. Heparin's anticoagulant effect can be rapidly reversed by the administration of protamine sulfate. However, this is usually not necessary, since reduction or omission of a heparin dose usually improves hemostasis and stops bleeding. Thrombocytopenia occurs in ~10% of recipients and is usually mild, with the platelet count falling to 50,000 to 100,000/uL. Thrombocytopenia is more common in patients receiving heparin derived from beef lung as opposed to porcine intestinal mucosa.LMWH is less likely to cause either thrombocytopenia or bleeding. However, antibodies arising from exposure to unfractionated heparin often crossreact with LMWH. Thus, LMWH cannot usually be used to treat patients with established thrombocytopenia.

Heparin-induced thrombocytopenia (HIT) results from generation of an autoantibody to a complex of unfractionated heparin with the anti-heparin protein platelet factor 4 (PF-4). Heparin-PF-4-antibody complexes can bind to the platelet Fc receptor and cause platelet activation, agglutination, and arterial thrombosis. Recognition of the rare complication of thrombocytopenia and paradoxical thrombosis is critical, since discontinuing heparin can promptly reverse the syndrome and may be lifesaving. Heparin administration for >5 months also carries a risk of osteoporosis, perhaps through its activation of osteoclasts.LMWHcauses less osteoporosis on chronic administration.

CHRONIC ORAL ANTICOAGULATION

The coumarin anticoagulants, which include warfarin and dicumarol (dicoumarol), prevent the reduction of vitamin K epoxides in liver microsomes and induce a state of vitamin K deficiency (see Fig. 117-1). They slow thrombin generation and clot formation by impairing the biologic activity of the prothrombin complex proteins and are used to prevent the recurrence of venous thrombosis and pulmonary embolism. Although regimens employing loading doses of drug have been advocated, the simplest way to induce anticoagulation is to administer a single dose of a coumarin compound and monitor the prothrombin time (PT) until the desired prolongation is achieved. For example, treatment can be initiated with 5 mg/d of warfarin or equivalent, with the goal of prolonging the PT to 1.5 to 2 times the control value. Although the PT may reach this value after a few days of therapy, effective anticoagulation, with stable reduction of all the prothrombin complex proteins, requires at least 1 week of warfarin administration. Most patients require a daily maintenance dose of 2.5 to 7.5 mg of warfarin to remain anticoagulated.

Because commercial thromboplastins have different potencies, thePT can vary widely.

In an effort to standardize oral anticoagulation, the International Normalized Ratio (INR) method has been adopted by most hospital laboratories and clinicians. In this reporting method, the ratio of the patient's PT is compared to the mean PT for a group of normal individuals. The ratio is adjusted for the sensitivity of the laboratory's thromboplastin determined by the International Sensitivity Index (ISI). Thus, INR $=(PT_{patient}/PT_{normal})_{ISI}$. Use of the INR permits physicians to obtain the appropriate level of anticoagulation independent of laboratory reagents and to follow published recommendations for intensity of anticoagulation. The intensity of anticoagulation may be varied somewhat depending on the clinical indication (Table 118-2). Patients with chronic indwelling venous catheters have been maintained on 1 mg/d of warfarin to prevent clot formation at the catheter tip; such a dose has no effect on the PT.

Although warfarin anticoagulants reduce the recurrence of deep venous thrombosis and pulmonary or cerebral embolism, they may also cause bleeding. Any patient who takes oral anticoagulants requires frequent monitoring of thePT. Despite the most careful management, fluctuations in PT can occur. Various drugs that alter liver microsomal metabolism of coumarins or compete for albumin-binding sites can increase or decrease the potency of these drugs (Table 118-3).

The risk of bleeding increases and, up to a point, the risk of recurrent thrombosis declines with the duration of anticoagulation. Patients with a single uncomplicated thromboembolic event achieve maximal benefit after 3 to 6 months of anticoagulation. About 10% of patients on an oral anticoagulant for 1 year have a bleeding complication requiring medical supervision, and 0.5 to 1% have a fatal hemorrhage. The anticoagulant effects of coumarins can be reversed by infusion of fresh-frozen plasma or by the administration of vitamin K. Fresh-frozen plasma works immediately, but the effects last only a few hours. Vitamin K takes 8 to 12 h to become effective; after vitamin K administration, vitamin K antagonists are more difficult to use for reinduction of anticoagulation. In many cases, reduction or omission of several doses of warfarin improves hemostasis and stops hemorrhage. Despite the risk of bleeding, some patients may require lifelong anticoagulation.

Hemorrhagic skin necrosis is a rare complication. Some patients with this complication are deficient in protein C, an anticoagulant protein whose activity is reduced by vitamin K antagonists. Patients suspected of protein C deficiency should only begin oral anticoagulant therapy when combined with heparin or plasma infusions to restore protein C levels to normal. Patients with an inherited coumarin resistance may require extremely high doses to get an anticoagulant effect. Psychologically disturbed patients may surreptitiously ingest coumarin and present with unexplained bleeding and a prolongedPT. Plasma coumarin levels can be measured to confirm such ingestion.

FIBRINOLYTIC THERAPY

Fibrinolysis, an important part of the normal hemostatic process, is initiated by the release of either tissue plasminogen activator (tPA) or pro-urokinase (proUK) from endothelial cells. These agents preferentially activate plasminogen adsorbed onto fibrin clots, a mechanism that localizes the lytic process to sites that contain fibrin thrombi. Although fibrinolysis begins immediately after vascular injury, clot lysis and vessel recanalization may not be complete for 7 to 10 days. The fibrinolytic pathway is

important in normal hemostasis; defects can predispose patients to either hemorrhage or recurrent thrombosis (Chap. 117). Activators of the fibrinolytic system are frequently used to accelerate clot lysis in patients with thromboembolism (Fig. 118-1;Table 118-4).

The pharmacologic agents being used to accelerate clot lysis are either derived from natural products or are chemically modified derivatives. They differ with respect to fibrin specificity and some types of complications (Table 118-4). For example, many individuals have antistreptococcal antibodies that react withSK and reduce its potency and cause febrile reactions. All fibrinolytic agents cause hemorrhage. In addition totPA andproUK, several other agents are relatively "fibrin-specific" and preferentially activate plasminogen in the presence of fibrin. Although this makes it theoretically possible to achieve selective clot lysis, in practice the efficacy and toxicity of the "specific" and "nonspecific" fibrinolytic agents are similar. However, equivalent doses ofrtPAcost 10 times more than SK.

Some systemic fibrinolysis always occurs after the infusion of clinically effective doses of fibrin-specific agents. Fibrinogen level falls ~25%after infusion of lytic doses ofrtPA. In addition, both the fibrin-specific and -nonspecific agents can cause hemorrhage as they cannot differentiate between vital hemostatic plugs and pathologic thrombi. To minimize the risk of bleeding, systemic lytic therapy is not recommended for patients with recent surgery or a history of neurologic lesions, gastrointestinal bleeding, or hypertension.

The current indications for fibrinolytic therapy are listed inTable 118-5. Fibrinolytic therapy is currently recommended for patients with massive pulmonary embolism that is complicated by hypotension, hypoxemia, and right heart strain. It is also used for selected patients with peripheral arterial embolism or occlusion and for patients with extensive iliofemoral thrombophlebitis. While lytic therapy may hasten the resolution of venous thrombi, the long-term benefit still remains unproven, and no firm evidence proves that lytic therapy reduces postphlebitic complications. In contrast, fibrinolytic therapy may be of distinct benefit in patients with axillary vein thrombosis, a condition that does not usually respond to conventional anticoagulation. Fibrinolytic agents are also used to restore the patency of occluded venous catheters and dialysis shunts. For this indication the agents are instilled locally. The extensive literature on the use of fibrinolytic agents to treat patients with coronary artery disease and myocardial infarction is reviewed inChap. 243. When given within a few hours of infarction, fibrinolytic therapy reduces mortality and myocardial damage.

Although the doses and mode of administration may differ slightly, the general principles and complications are the same for all the fibrinolytic agents.SK andUK are the oldest and most extensively studied fibrinolytic agents. SK is a bacterial enzyme, and UK is a product of renal tubular epithelial cells. SK is an indirect plasminogen activator that interacts with circulating plasminogen to form an equimolar complex with proteolytic activity. The SK-plasminogen complex then activates additional plasminogen molecules that initiate fibrinolysis. In contrast, UK has intrinsic proteolytic activity and can activate plasminogen directly.

In the case ofSK, a loading dose of 250,000 units is usually given irrespective of body weight. Since patients may have antistreptococcal antibodies, the loading dose may need to be repeated. In addition, patients may develop acute allergic symptoms

including urticaria and, occasionally, serum sickness reactions. WithUK, a loading dose of 4400 units per kilogram body weight is administered over 10 to 30 min. Both regimens induce an intense lytic state as evidenced by a drop in fibrinogen, prolongation of the thrombin time, and a prolongation of the euglobulin lysis time -- an in vitro measure of fibrinolytic activity. After the initial loading dose, 100,000 units of SK or 4400 units of UK per kilogram body weight are administered hourly for 24 to 72 h. At the desired time, the lytic state is reversed by discontinuing UK or SK and by administering heparin for 7 to 10 days. Heparin can be started at the same time as the fibrinolytic agent. Fibrinolytic therapy should be initiated as soon as possible after the onset of thrombosis or embolism.

Fibrin-specific agents such asrtPA orproUKare also administered intravenously. Systemic infusion of 100 mg rtPA over 6 h restores coronary artery patency in ~75% of patients. Patients are then maintained on heparin for several days. ProUK given in a similar manner has almost identical effects, but large clinical trials suggest that rtPA is superior to other fibrinolytic agents in maintaining patency of acutely occluded coronary arteries.

ANTIPLATELET DRUG THERAPY

Antiplatelet drugs play a critical role in the management of patients with arterial vascular disease and thromboembolism. Aspirin is the most widely studied of these drugs because of its unique pharmacology. A single dose of aspirin irreversibly acetylates and inactivates the enzyme cyclooxygenase and thereby inhibits platelet production of thromboxane A_2. Although aspirin may also inactivate cyclooxygenase in other tissues, including endothelial cells, such cells recover rapidly by synthesizing new enzyme. Platelets, which are anucleate, cannot synthesize new enzyme and remain inactive for the rest of their life span. As little as one 160-mg tablet of aspirin daily or a 325-mg tablet every other day inhibits platelet thromboxane production and aggregation.

Patients with coronary artery disease who have unstable angina are at high risk for myocardial infarction (Chap. 243). The prompt administration of aspirin dramatically reduces progression to myocardial infarction in this group, although aspirin has no effect on the frequency, intensity, or duration of chronic angina. Aspirin also reduces the incidence of second infarction by 25% when administered to patients who have had a myocardial infarct. Daily aspirin therapy also reduces the incidence of first infarcts. The combination of aspirin and dipyridamole, when begun before surgery, may also increase the patency of coronary bypass grafts; the same combination reduces the incidence of cerebral emboli in patients on warfarin who have prosthetic intracardiac valves. Although dipyridamole has been a popular antithrombotic agent, it has little efficacy when given alone. Aspirin may be the active agent in the combination aspirin-dipyridamole trials.

Aspirin also reduces the frequency of transient ischemic attacks in patients with occlusive cerebrovascular disease. It has largely supplanted anticoagulation with the coumarin compounds in patients with transient ischemia. Aspirin also reduces the incidence of a second stroke by 25% when administered to men following a first stroke. Aspirin is also effective in maintaining the patency of arteriovenous cannulas inserted into patients with renal failure who require hemodialysis. Aspirin plus dipyridamole may

also slow the progression of some forms of glomerulonephritis, although these drugs are not widely used in the treatment of renal disease. Aspirin is not effective in maintaining the patency of vessels following percutaneous angioplasty or stent placement.

Although aspirin is clearly the most efficacious antiplatelet agent in clinical use today, a large number of new drugs are being tested that may supplement aspirin therapy. Ticlopidine, a potent inhibitor of platelet function, is effective as an alternative to aspirin in patients with cerebrovascular disease and is superior to aspirin or warfarin in maintaining coronary stent patency. Ticlopidine is more expensive than aspirin and causes some serious side effects, including neutropenia and occasional rare episodes of thrombotic thrombocytopenic purpura (TTP). A related drug, clopidogrel (Plavix), has been proposed as a substitute, but rare cases of TTP have also been noted with its use.

Monoclonal antibodies and both recombinant and chemically synthesized peptides that block either platelet adhesion or aggregation are being tested in clinical trials. A monoclonal antibody Fab fragment that blocks fibrinogen binding to platelet GpIIb/IIIa, thus inhibiting platelet aggregation (abciximab, RheoPro), is used in patients with coronary artery disease who undergo angioplasty. RheoPro is also being evaluated in other settings, e.g., as an adjunct to fibrinolytic therapy in patients with an acute myocardial infarction. A cyclic peptide based on the consensus fibronectin adhesion sequence RGD (arginine, glycine, aspartic acid) called eptifibatide (Integrilin) is as effective as RheoPro for maintaining patency after angioplasty and stent placement as is a small molecule inhibitor (Aggrestat). Orally active GpIIb/IIIa inhibitors have not been proven safe or effective.

The uses of antithrombotic therapy are evolving rapidly. However, aspirin is the current mainstay for chronic therapy and should be used (325 qd or qod) indefinitely in any patient who has had a coronary or cerebral thrombosis. It will reduce ischemic events by 25% or more. Patients undergoing angioplasty or stent placement should receive RheoPro/Integrilin or Aggrestat acutely, followed by 3 weeks of ticlopidine or clopidogrel.

(Bibliography omitted in Palm version)

Back to Table of Contents

PART SEVEN -INFECTIOUS DISEASES

SECTION 1 -BASIC CONSIDERATIONS IN INFECTIOUS DISEASES

119. INTRODUCTION TO INFECTIOUS DISEASES: HOST-PARASITE INTERACTIONS - *Lawrence C. Madoff*, *Dennis L. Kasper*

Despite decades of dramatic progress in their treatment and prevention, infectious diseases remain a major cause of death and debility and are responsible for worsening the living conditions of many millions of people around the world. Infections frequently challenge the physician's diagnostic skill and must be considered in the differential diagnoses of syndromes affecting every organ system.

CHANGING EPIDEMIOLOGY OF INFECTIOUS DISEASES

With the advent of antimicrobial agents, some medical leaders believed that infectious diseases would soon be eliminated and become of historic interest only. Indeed, the hundreds of chemotherapeutic agents developed since World War II, most of which are potent and safe, include drugs effective not only against bacteria but also against viruses, fungi, and parasites. Nevertheless, we now realize that as we developed antimicrobial agents, microbes developed the ability to elude our best weapons and to counterattack with new survival strategies. Antibiotic resistance occurs at an alarming rate among all classes of mammalian pathogens. Pneumococci resistant to penicillin and enterococci resistant to vancomycin have become commonplace. Even *Staphylococcus aureus* with reduced susceptibility to vancomycin has appeared. Such pathogens present real clinical problems in managing infections that were easily treatable just a few years ago. Diseases once thought to have been nearly eradicated from the developed world -- tuberculosis, cholera, and rheumatic fever, for example -- have rebounded with renewed ferocity. Newly discovered and emerging infectious agents appear to have been brought into contact with humans by changes in the environment and by movements of human and animal populations. An example of the propensity for pathogens to escape from their usual niche is the alarming 1999 outbreak in New York of encephalitis due to a flavivirus similar or identical to West Nile virus, which had never previously been isolated in the Americas.

Many infectious agents have been discovered only in recent decades. Ebola virus, hantavirus, the agent of human granulocytotropic ehrlichiosis, and retroviruses such as HIV humble us despite our deepening understanding of pathogenesis at the most basic molecular level. Even in developed countries, infectious diseases have made a resurgence. Between 1980 and 1996, mortality from infectious diseases in the United States increased by 64% to levels not seen since the 1940s.

The role of infectious agents in the etiology of diseases once believed to be noninfectious is being increasingly recognized. For example, it is now widely accepted that *Helicobacter pylori* is the causative agent of peptic ulcer disease and perhaps of gastric malignancy. Human papillomavirus is likely to be the most important cause of invasive cervical cancer. A new human herpesvirus (HHV-8) is believed to be the cause of most cases of Kaposi's sarcoma. Epstein-Barr virus is a cause of certain lymphomas and may play a role in the genesis of Hodgkin's disease. The possibility certainly exists

that other diseases of unknown cause, such as rheumatoid arthritis, sarcoidosis, or inflammatory bowel disease, have infectious etiologies. There is even evidence that atherosclerosis may have an infectious component.

Medical advances over infectious diseases have been hindered by changes in the patient population. Immunocompromised hosts now constitute a significant proportion of the seriously infected population. Physicians immunosuppress their patients to prevent the rejection of transplants and to treat neoplastic and inflammatory diseases. Some infections, most notably that caused by HIV, immunocompromise the host in and of themselves. Lesser degrees of immunosuppression are associated with other infections, such as influenza and syphilis. Infectious agents that coexist peacefully with immunocompetent hosts wreak havoc in those who lack a complete immune system. AIDS has brought to prominence once-obscure organisms such as *Pneumocystis carinii*, *Cryptosporidium parvum*, and *Mycobacterium avium*.

BIOTERRORISM

In recent years, the efforts of some governments and terrorist organizations to use biological weaponry have refocused public concern on the topic. To date, there is little evidence that biological weapons have ever been effectively used; indeed, their ease of use may be overstated. However, the ability of infectious agents to inflict widespread illness and thus to cause societal disruption and panic, together with the relatively low cost of these agents, has led to their being called a "poor man's nuclear arsenal."

Several pathogens have been considered likely candidates for biological warfare. *Bacillus anthracis*, which causes the zoonosis anthrax, is widely viewed as the leading contender. The hardy spores of the bacillus can be distributed by bombardment or other dispersal mechanisms. Inhalation of this pathogen results in severe pneumonia with a mortality rate of 95% in untreated persons. A World Health Organization report estimated that 50 kg of *B. anthracis* released upwind of a city with a population of 500,000 would result in up to 95,000 fatalities, with an additional 125,000 persons incapacitated. Moreover, the attack might go undetected until large numbers of seriously ill individuals presented with overt disease. In 1979, an accidental release from a military microbiology facility near Sverdlovsk in the former Soviet Union caused at least 66 deaths from inhalational anthrax along a 4-km-wide path downwind of the facility. The vast scope of the Soviet Union's biological warfare program, employing 60,000 people at its height, has only recently come to light. This endeavor is thought to be echoed by efforts in many other countries in the world today. In response to the perceived threat of anthrax as a biological weapon, the U.S. military recently decided to vaccinate more than 2 million of its members against this infection.

Smallpox, an ancient scourge caused by variola virus, has also been considered as a bioweapon owing to its contagiousness and high mortality rate and to the declining population of immunized persons. Indeed, one of the earliest accounts of biological warfare involved the distribution of smallpox-infected blankets to Native American tribes by British troops. Debate continues about whether to eradicate the two known existing stocks of the virus in U.S. and Russian government laboratories. Many investigators believe that additional undocumented stockpiles of the virus exist around the world.

Other infectious organisms that combine the virulence and stability necessary for biological weapons include *Yersinia pestis*, the agent of plague, and *Francisella tularensis*, the agent of tularemia. Viral hemorrhagic fever agents such as the Ebola and Marburg viruses as well as toxins such as that from *Clostridium botulinum* have also been considered as biological weapons. While some of the diseases caused by these agents can be effectively treated or prevented if sufficient resources exist to do so, it may also be possible for an aggressor to render organisms resistant to antibiotics or even to vaccines through genetic manipulation or other means.

HOST FACTORS IN INFECTION

For any infectious process to occur, the parasite and the host must first encounter each other. Factors such as geography, environment, and behavior thus influence the likelihood of infection. Though the initial encounter between a susceptible host and a virulent organism frequently results in disease, some organisms can be harbored in the host for years before disease becomes clinically evident. For a complete view, individual patients must be considered in the context of the population to which they belong. Infectious diseases do not often occur in isolation; rather, they spread through a group exposed from a point source (e.g., a contaminated water supply) or from individual to individual (e.g., via respiratory droplets). Thus, the clinician must be alert to infections prevalent in the community as a whole. A detailed history, including information on travel, behavioral factors, exposures to animals or potentially contaminated environments, and living and occupational conditions, must be elicited. For example, the likelihood of infection by *Plasmodium falciparum* can be significantly affected by altitude, climate, terrain, season, and even time of day. Antibiotic-resistant strains are localized to specific geographic regions, and a seemingly minor alteration in a travel itinerary can dramatically influence the likelihood of acquiring chloroquine-resistant malaria. If such important details in the history are overlooked, inappropriate treatment may result in the death of the patient. Likewise, the chance of acquiring a sexually transmitted disease can be greatly affected by a relatively minor variation in sexual practices, such as the method used for birth control. Knowledge of the relationship between specific risk factors and disease allows the physician to influence a patient's health even before the development of infection by modification of these risk factors and -- when a vaccine is available -- by immunization.

Many specific host factors influence the likelihood of acquiring an infectious disease. Age, immunization history, prior illnesses, level of nutrition, pregnancy, coexisting illness, and perhaps emotional state all have some impact on the risk of infection after exposure to a potential pathogen. The importance of individual host defense mechanisms, either specific or nonspecific, becomes apparent in their absence, and our understanding of these immune mechanisms is enhanced by studies of clinical syndromes developing in immunodeficient patients (Table 119-1). For example, the frequent occurrence of meningococcal disease in people with deficiencies in specific complement proteins of the "membrane attack complex" underscores the importance of an intact complement system in the prevention of meningococcal infection.

Medical care itself increases the patient's risk of acquiring an infection in several ways: (1) through contact with pathogens during hospitalization, (2) through breaching of the skin (with intravenous devices or surgical incisions) or mucosal surfaces (with

endotracheal tubes or bladder catheters), (3) through introduction of foreign bodies, (4) through alteration of the natural flora with antibiotics, and (5) through treatment with immunosuppressive drugs.

THE IMMUNE RESPONSE

Infection involves complicated interactions of parasite and host and inevitably affects both. In most cases, a pathogenic process consisting of several steps is required for the development of infections. Since the competent host has a complex series of barricades in place to prevent infection, the successful parasite must utilize specific strategies at each of these steps. The specific strategies used by bacteria, viruses, and parasites (Chaps. 120 and180) have some remarkable conceptual similarities, but the strategic details are unique not only for each class of organism but also for individual species within a class.

Once in the bloodstream or a normally sterile body site, the microorganism faces the host's tightly integrated cellular and humoral immune systems. Cellular immunity (Chap. 305), comprising T lymphocytes, macrophages, and natural killer cells, primarily recognizes and combats pathogens that proliferate intracellularly. Cellular immune mechanisms are important in immunity to all classes of infectious agents, including most viruses and many bacteria (e.g., *Mycoplasma*, *Chlamydia*, *Listeria*, *Salmonella*, and *Mycobacterium*), parasites (e.g., *Trypanosoma*, *Toxoplasma*, and *Leishmania*), and fungi (e.g., *Histoplasma*, *Cryptococcus*, and *Coccidioides*). Usually, T lymphocytes are activated by macrophages and B lymphocytes, which present foreign antigens along with the host's own major histocompatibility complex antigen. Activated T cells may then act in several ways to fight infection. Cytotoxic T cells may directly attack and lyse host cells that express foreign antigens. Helper T cells stimulate the proliferation of B cells and the production of immunoglobulins. B cells and T cells communicate with each other via a variety of signals, and often more than one signal is employed simultaneously. For example, costimulation through the CD40-CD40 ligand increases B cell responses, and costimulation via the B7-CD28 axis is required for activation of the CD4+ helper T cell. T cells elaborate cytokines (e.g., interferon), which directly inhibit the growth of pathogens or stimulate killing by host macrophages and cytotoxic cells. Cytokines also augment the host's immunity by stimulating the inflammatory response (fever, the production of acute-phase serum components, and the proliferation of leukocytes). Cytokine stimulation does not always result in a favorable response in the host; septic shock (Chap. 124) and toxic shock syndrome (Chaps. 139 and140) are among the conditions that are mediated by these inflammatory substances.

The reticuloendothelial system comprises monocyte-derived phagocytic cells that are located in the liver (Kupffer cells), lung (alveolar macrophages), spleen, kidney (mesangial cells), brain (microglia), and lymph nodes and that clear circulating microorganisms. Although these tissue macrophages and polymorphonuclear leukocytes (PMNs) are capable of killing microorganisms without help, they function much more efficiently when pathogens are first *opsonized* (Greek, "to prepare for eating") by components of the complement system such as C3b and/or by antibodies.

Extracellular pathogens, including most encapsulated bacteria, are attacked by the humoral immune system, which includes antibodies, the complement cascade, and

phagocytic cells. Antibodies are complex glycoproteins (also called immunoglobulins) that are produced by mature B lymphocytes, circulate in body fluids, and are secreted on mucosal surfaces. Antibodies specifically recognize and bind to foreign antigens. One of the most impressive features of the immune system is the ability to generate an incredible diversity of antibodies capable of recognizing virtually every foreign antigen yet not reacting with self. In addition to being exquisitely specific for antigens, antibodies come in different structural and functional classes: IgG predominates in the circulation and persists for many years after exposure; IgM is the earliest specific antibody to appear in response to infection; secretory IgA is important in immunity at mucosal surfaces, while monomeric IgA appears in the serum; and IgE is important in allergic and parasitic diseases. Antibodies may directly impede the function of an invading organism, neutralize secreted toxins and enzymes, or facilitate the removal of the antigen (invading organism) by phagocytic cells. Immunoglobulins participate in cell-mediated immunity by promoting the antibody-dependent cellular cytotoxicity functions of certain T lymphocytes. Antibodies also promote the deposition of complement components on the surface of the invader.

The complement system (Chap. 305) consists of a group of serum proteins functioning as a cooperative, self-regulating cascade of enzymes that adhere to -- and in some cases disrupt -- the surface of invading organisms. Some of these surface-adherent proteins (e.g., C3b) can then act as opsonins for destruction of microbes by phagocytes. The later, "terminal" components (C7, C8, and C9) can directly kill some bacterial invaders (notably, many of the neisseriae) by forming a "membrane attack complex" and disrupting the integrity of the bacterial membrane, thus causing bacteriolysis. Other complement components, such as C5a, act as chemoattractants forPMNs. Complement activation and deposition occur by either or both of two pathways: the classic pathway is activated primarily by immune complexes (i.e., antibody bound to antigen), and the alternative pathway is activated by microbial components, frequently in the absence of antibody. PMNs have receptors for both antibody and C3b, and antibody and complement function together to aid in the clearance of infectious agents.

PMNs, short-lived white blood cells that engulf and kill invading microbes, are first attracted to inflammatory sites by chemoattractants such as C5a, which is a product of complement activation at the site of infection. PMNs localize to the site of infection by adhering to cellular adhesion molecules expressed by endothelial cells. Endothelial cells express these receptors, called *selectins* (CD-62, ELAM-1), in response to inflammatory cytokines such as tumor necrosis factor (TNF) a and interleukin 1. The binding of these selectin molecules to specific receptors on PMNs results in the adherence of the PMNs to the endothelium. Cytokine-mediated upregulation and expression of intercellular adhesion molecule 1 (ICAM 1) on endothelial cells then take place, and this latter receptor binds to$b_2$integrins on PMNs, thereby facilitating diapedesis into the extravascular compartment. Once the PMNs are in the extravascular compartment, various molecules such as arachidonic acids further enhance the inflammatory process.

Approach to the Patient

The clinical manifestations of infectious diseases at presentation are myriad, varying from fulminant life-threatening processes to brief and self-limited conditions to indolent chronic maladies. The clinician must use all the skills of medicine to diagnose the

infection and prescribe appropriate treatment. First, a careful history is essential and must include details on underlying chronic diseases; medications; occupation; travel; and risk factors for exposure to certain types of pathogens, such as those associated with sexual contacts, family illnesses, illicit drug use, particular animals, blood transfusions, ingestion of contaminated liquids or foods, or bites of insect vectors. Since infectious diseases may involve many organ systems, a careful review of systems may elicit important clues as to the disease process. The physical examination must be thorough, and attention must be paid to seemingly minor details: a soft heart murmur that might indicate bacterial endocarditis; an evanescent skin rash that suggests rheumatic fever; or a retinal lesion that suggests disseminated candidiasis or cytomegalovirus (CMV) infection.

LABORATORY INVESTIGATIONS

Laboratory studies must be carefully considered and directed toward establishing an etiologic diagnosis in the shortest possible time, at the lowest possible cost, and with the least possible discomfort to the patient. Cultures must be performed in a manner that minimizes the likelihood of contamination with normal flora while maximizing the yield. A sputum sample is far more likely to be valuable when elicited with careful coaching by the clinician than when collected in a container simply left at the bedside with cursory instructions. Gram's stains of specimens should be interpreted carefully and the quality of the specimen assessed. The findings on Gram's staining should correspond to the results of culture; a discrepancy may suggest diagnostic possibilities such as infection due to fastidious or anaerobic bacteria.

The microbiology laboratory must be an ally in the diagnostic endeavor (Chap. 121). Astute laboratory personnel will suggest optimal culture and transport conditions or alternative tests to facilitate diagnosis. If informed about specific potential pathogens, an alert laboratory staff will allow sufficient time for these organisms to become evident in culture, even when present in small numbers or when slow-growing. The parasitology technician who is attuned to the specific diagnostic considerations relevant to a particular case may be able to detect the rare, otherwise-elusive egg or cyst in a stool specimen. In cases where a diagnosis appears difficult, serum should be stored during the early acute phase of the illness so that a diagnostic rise in titer of antibody to a specific pathogen can be detected later. Bacterial and fungal antigens can sometimes be detected in body fluids, even when cultures are negative or are rendered sterile by antibiotic therapy. Techniques such as the polymerase chain reaction allow the amplification of specific DNA sequences so that minute quantities of foreign nucleic acids can be recognized in host specimens.

TREATMENT

Optimal therapy for infectious diseases requires a broad knowledge of medicine and careful clinical judgment. Life-threatening infections such as bacterial meningitis or sepsis, viral encephalitis, or falciparum malaria must be treated immediately, often before a specific causative organism is identified. Antimicrobial agents must be chosen empirically and must be active against the range of potential infectious agents consistent with the clinical scenario. In contrast, good clinical judgment sometimes dictates withholding of antimicrobials in a self-limited process or until a specific

diagnosis is made. The dictum *primum non nocere* should be adhered to, and it should be remembered that all antimicrobials carry a risk (and a cost) to the patient. Direct toxicity may be encountered -- e.g., ototoxicity due to aminoglycosides, lipodystrophy due to HIV protease inhibitors, and hepatotoxicity due to antituberculous agents such as isoniazid and rifampin. Allergic reactions are common and can be serious. Since superinfection sometimes follows the eradication of the normal flora and colonization by a resistant organism, one invariable principle is that infectious disease therapy should be directed toward as narrow a spectrum of infectious agents as possible. Treatment specific for the pathogen should result in as little perturbation as possible of the host's microflora. With few exceptions, abscesses require surgical or percutaneous drainage for cure. Foreign bodies, including medical devices, must generally be removed in order to eliminate an infection of the device or of the adjacent tissue. Other infections, such as necrotizing fasciitis, peritonitis due to a perforated organ, gas gangrene, and chronic osteomyelitis, require surgery as the primary means of cure; in these conditions, antibiotics play only an adjunctive role.

The role of immunomodulators in the management of infectious diseases has received increasing attention. Glucocorticoids have been shown to be of benefit in the treatment of *Haemophilus influenzae* meningitis in children and in therapy for *P. carinii* pneumonia in patients with AIDS. The use of these agents in other infectious processes remains less clear and in some cases (in cerebral malaria and septic shock, for example) is detrimental. Other agents that modulate the immune response include prostaglandin inhibitors, specific lymphokines, andTNFinhibitors. Specific antibody therapy plays a role in the treatment and prevention of many diseases. Specific immunoglobulins have long been known to prevent the development of symptomatic rabies and tetanus. More recently,CMVimmune globulin has been recognized as important not only in preventing the transmission of the virus during organ transplantation but also in treating CMV pneumonia in bone marrow transplant recipients. There is a strong need for well-designed clinical trials to evaluate each new interventional modality.

PERSPECTIVE

The genetic simplicity of many infectious agents allows them to undergo rapid evolution and to develop selective advantages that result in constant variation in the clinical manifestations of infection. Moreover, changes in the environment and the host can predispose new populations to a particular infection. An epidemic of lethal respiratory failure -- later identified as hantavirus pulmonary syndrome -- on a Navajo reservation in the southwestern United States in 1993 caused nationwide alarm, exemplifying the fear that new plagues induce in the human psyche.

The potential for infectious agents to emerge in novel and unexpected ways requires that physicians and public health officials be knowledgeable, vigilant, and open-minded in their approach to unexplained illness. The emergence of antimicrobial-resistant pathogens (e.g., enterococci that are resistant to all known antimicrobial agents and cause infections that are essentially untreatable) has led some to conclude that we are entering the "postantibiotic era." Others have held to the perception that infectious diseases no longer represent as serious a concern to world health as they once did. The progress that science, medicine, and society as a whole have made in combating these maladies is impressive, and it is ironic that, as we stand on the threshold of an

understanding of the most basic biology of the microbe, infectious diseases are posing renewed problems. We are threatened by the appearance of new diseases such as AIDS, hepatitis C, and Ebola virus infection and by the reemergence of old foes such as tuberculosis, cholera, plague, and *Streptococcus pyogenes* infection. True students of infectious diseases were perhaps less surprised than anyone else by these developments. Those who know pathogens are aware of their incredible adaptability and diversity. As ingenious and successful as therapeutic approaches may be, our ability to develop methods to counter infectious agents so far has not matched the myriad strategies employed by the sea of microbes that surrounds us. Their sheer numbers and the rate at which they can evolve are daunting. Moreover, environmental changes, rapid global travel, population movements, and medicine itself -- through its use of antibiotics and immunosuppressive agents -- all increase the impact of infectious diseases. Although new vaccines, new antibiotics, improved global communication, and new modalities for treating and preventing infection will be developed, pathogenic microbes will continue to develop new strategies of their own, presenting us with an unending and dynamic challenge.

(Bibliography omitted in Palm version)

Back to Table of Contents

120. MOLECULAR MECHANISMS OF MICROBIAL PATHOGENESIS- *Gerald B. Pier*

Over the past two decades, molecular studies of microbial pathogenesis have yielded an explosion of information about the various microbial and host molecules that contribute to the processes of infection and disease. These processes can be classified into several stages: microbial encounter with and entry into the host; microbial growth after entry; avoidance of innate host defenses; tissue invasion and tropism; tissue damage; and transmission to new hosts. Virulence is the measure of an organism's capacity to cause disease and is a function of the pathogenic factors elaborated by microbes. These factors promote colonization (the simple presence of potentially pathogenic microbes in or on a host), infection (attachment and growth of pathogens and avoidance of host defenses), and disease (often, but not always, the activities of secreted toxins or toxic metabolites). In addition, the host's inflammatory response to infection greatly contributes to disease and its attendant clinical signs and symptoms. Knowledge of the molecular structures of the microbial surface, their interactions with the host, and the host response is critical to an understanding of the basic processes of infection and disease.

MICROBIAL ENTRY AND ADHERENCE

ENTRY SITES

A microbial pathogen can potentially enter any part of a host organism. In general, the type of disease produced by a particular microbe is often a direct consequence of its route of entry into the body. The most common sites of entry are body parts in contact with the external environment, including mucosal surfaces (particularly those of the respiratory, alimentary, and urogenital tracts) and the skin. Ingestion, inhalation, and sexual contact are typical routes of microbial entry. Other portals of entry include injuries to the skin (cuts, bites, burns, trauma) along with injection via natural (i.e., vector-borne) or artificial (i.e., needle-stick injury) routes. A few pathogens, such as *Schistosoma* spp., can penetrate unbroken skin. The conjunctiva can serve as an entry point for pathogens of the eye.

Microbial entry usually relies on the organism's biologic characteristics and reflects the presence of specific microbial factors needed for persistence and growth in a tissue. Fecal-oral spread via the alimentary tract requires a biology consistent with survival in the varied environments of the gastrointestinal tract (including the low pH of the stomach and the high bile content of the intestine) as well as in contaminated food or water outside the host. Organisms that gain entry via the respiratory tract are most often those that survive well in small moist droplets produced during sneezing and coughing; most such pathogens do not survive well once they dry out. Pathogens that enter by venereal routes often survive best on the warm moist environment of the urogenital mucosa. Many sexually transmitted human pathogens have restricted host ranges and do not infect other animals (e.g., *Neisseria gonorrhoeae*, *Treponema pallidum*, and HIV).

The biology of microbes entering through the skin is highly varied. Some of these organisms can survive in a broad range of environments, such as the salivary glands or alimentary tracts of arthropod vectors, the mouths of larger animals, soil, and water. A

complex biology allows protozoan parasites such as *Plasmodium*, *Leishmania*, and *Trypanosoma* spp. to undergo morphogenic changes that allow the organism to be transmitted to mammalian hosts during insect feeding for blood meals. Plasmodia are injected as infective sporozoites from the salivary glands during mosquito feeding. *Leishmania* parasites are regurgitated as promastigotes from the alimentary tract of sandflies and injected by bite into a susceptible host. Trypanosomes are first ingested from infected hosts by reduviid bugs; the pathogens then multiply in the gastrointestinal tract of the insects and are released in feces onto the host's skin during subsequent feedings. Most microbes that land directly on intact skin are destined to die, as survival on the skin or in hair follicles requires resistance to fatty acids, low pH, and other antimicrobial factors on skin. Once it is damaged (and particularly if it becomes necrotic), the skin can be a major portal of growth and entry for pathogens or their toxic products. Tetanus and burn wound infections are clear examples. After animal bites, pathogens resident in the animal's saliva gain access through the skin to the victim's tissues. Rabies is the paradigm for this pathogenic process; rabies virus grows in striated muscle cells at the site of inoculation.

MICROBIAL ADHERENCE

Once in or on a host, most microbes must anchor themselves to a tissue or tissue factor; the possible exceptions are organisms that directly enter the bloodstream and multiply there. Specific microbial ligands or adhesins for host receptors constitute a major area of study in the field of microbial pathogenesis. Adhesins comprise a wide range of surface structures, not only anchoring the microbe to a tissue and promoting cellular entry where appropriate but also eliciting host responses critical to the pathogenic process (Table 120-1). Most microbes produce multiple adhesins specific for multiple host receptors. These adhesins are often redundant, are serologically variable, and act additively or synergistically with other microbial factors to promote microbial sticking to host tissues. In addition, some microbes (such as *Mycobacterium tuberculosis* and *Legionella pneumophila*) adsorb host proteins (such as complement components) onto their surface and utilize the natural host protein receptor for microbial binding and entry into target cells.

All viral pathogens must bind to host cells, enter them, and replicate within them. Viral coat proteins serve as the ligands for cellular entry, and more than one ligand-receptor interaction may be needed; for example, HIV utilizes its envelope glycoprotein (gp) 120 to enter host cells by binding to both CD4 and one of several receptors for chemokines. Similarly, the measles virus H glycoprotein binds to both CD46 and the membrane-organizing protein moesin. The gC protein on herpes simplex virus binds to heparin sulfate; this step is followed by attachment to cells mediated by the viral gD (and possibly gH) protein. CD46 has now been shown to be the cellular receptor for human herpesvirus type 6. Eukaryotic parasites use complicated surface glycoproteins as adhesins, some of which are lectins with specificity for carbohydrates on host cells.

Among the microbial adhesins studied in greatest detail are bacterial pili and flagella. *Pili* or *fimbriae* are commonly used by gram-negative bacteria for attachment to host cells and tissues. In electron micrographs, these hairlike projections (up to several hundred per cell) may be confined to one end of the organism (polar pili) or distributed more evenly over the surface. An individual cell may have pili with a variety of functions.

Most pili are made up of a major pilin protein subunit (molecular weight, 17,000 to 30,000) that polymerizes to form the pilus. Many strains of *Escherichia coli* express mucus-binding type 1 pili, whose binding to host tissues is inhibited by D-mannose. Other strains produce the Pap (pyelonephritis-associated) or P pilus adhesin that mediates binding to digalactose (gal-gal) residues on globosides of the human P blood groups. These pili have proteins located at the tips of the main pilus unit that are critical to the binding specificity of the whole pilus unit. Immunization with the mannose-binding FimH tip protein of type 1 pili prevents experimental *E. coli* bladder infections in mice and monkeys. *E. coli* cells causing diarrheal disease express pilus-like receptors for enterocytes on the small bowel, along with other receptors termed *colonization factors*.

A common type of pilus found in *Neisseria* spp., *Moraxella* spp., *Vibrio cholerae*, and *Pseudomonas aeruginosa* mediates adherence of these organisms to target surfaces. These pili tend to have a relatively conserved amino-terminal region and a more variable carboxyl-terminal region. For some species such as *N. gonorrhoeae* and *Neisseria meningitidis*, the pili are critical for attachment to mucosal epithelial cells. For others, such as *P. aeruginosa*, the pili only partially mediate the cells' adherence to host tissues. *V. cholerae* cells appear to use two different types of pili for intestinal colonization. Whereas interference with this stage of colonization would appear to be an effective antibacterial strategy, attempts to develop pilus-based vaccines for human diseases have not been highly successful to date.

Flagella are long appendages attached at either one or both ends of the bacterial cell (polar flagella) or distributed over the entire cell surface (peritrichous flagella). Flagella, like pili, are composed of a polymerized or aggregated basic protein. In flagella, the protein subunits form a tight helical structure and vary serologically with the species. Spirochetes such as *T. pallidum* and *Borrelia burgdorferi* have axial filaments similar to flagella running down the long axis of the center of the cell, and they "swim" by rotation around these filaments. Some bacteria can glide over a surface in the absence of obvious motility structures.

Other bacterial structures involved in adherence to host tissues include specific staphylococcal and streptococcal proteins that bind to human extracellular matrix proteins such as fibrin, fibronectin, laminin, and collagen. Fibronectin appears to be a commonly used receptor for various pathogens; a particular sequence, Arg-Gly-Asp or RGD, is critical for bacterial binding. Surface lipoteichoic acids may also promote streptococcal adherence to mucosal surfaces. The surface lipopolysaccharide (LPS) of *P. aeruginosa* mediates binding to the cystic fibrosis transmembrane conductance regulator (CFTR) on airway epithelial cells. Coagulase-negative staphylococci readily colonize prosthetic devices and catheters commonly used in medical care; the surface capsular polysaccharide of these organisms promotes binding to the prosthetic material. It has been reported that *Staphylococcus aureus* produces the same capsular polysaccharide and may also use this material to colonize prosthetic devices.

FUNGAL ADHESINS

Several fungal adhesins have been described that mediate colonization of epithelial surfaces, particularly adherence to structures like fibronectin, laminin, and collagen. The product of the *Candida albicans INT1* gene, int1p, bears similarity to mammalian

integrins that bind to extracellular matrix proteins. Transformation of normally nonadherent *Saccharomyces cerevisiae* with this gene allows these yeast cells to adhere to human epithelial cells. Disruption of *INT1* in *C. albicans* diminishes but does not eliminate epithelial cell adhesion; this result indicates that both int1p and other adhesins mediate binding of *C. albicans* to epithelial cells. Moreover, int1p is needed for filamentous growth of *C. albicans* -- a phenotype linked to virulence, and particularly to the ability to penetrate keratinized epithelium. *INT1*-deficient *C. albicans* exhibits markedly reduced virulence in a mouse model of infection.

For several fungal pathogens that initiate infections after inhalation of infectious material, the inoculum is ingested by alveolar macrophages, in which the cells transform to pathogenic phenotypes. Like *C. albicans*, *Blastomyces dermatitidis* binds to CD11b/CD18 integrins as well as to CD14 on macrophages. *B. dermatitidis* produces a 120-kDa surface protein, designated WI-1, that mediates this adherence. The binding domain of WI-1 is homologous to the invasin protein of *Yersinia* that binds to the same type of host cell receptor. An unidentified factor on *Histoplasma capsulatum* also mediates binding of this fungal pathogen to the integrin surface proteins.

HOST RECEPTORS

Host receptors are found both on target cells (such as epithelial cells lining mucosal surfaces) and within the mucus layer covering these cells. Microbial pathogens bind to a wide range of host receptors to establish infection (Table 120-1). Selective loss of host receptors for a pathogen may confer natural resistance to an otherwise susceptible population. For example, *Plasmodium vivax*, one of four *Plasmodium* species causing malaria, binds to the Duffy blood group antigen, Fy, on erythrocytes. In West Africa, 70% of individuals lack Fy antigens and are resistant to *P. vivax* infection. *Salmonella typhi*, the etiologic agent of typhoid fever, usesCFTR to enter the gastrointestinal submucosa after being ingested. As homozygous mutations in *CFTR* are the cause of the life-shortening disease cystic fibrosis, heterozygote carriers (e.g., 4 to 5% of individuals of European ancestry) may have had a selective advantage due to decreased susceptibility to *S. typhi* infection.

Numerous virus-target cell interactions have been described, and it is now clear that different viruses can use similar host cell receptors for entry. The list of certain and likely host receptors for viral pathogens is long. Among the host membrane components that can serve as receptors for viruses are sialic acids, gangliosides, glycosaminoglycans, integrins and other members of the immunoglobulin superfamily, histocompatibility antigens, and regulators and receptors for complement components.

MICROBIAL GROWTH AFTER ENTRY

Once established on a mucosal or skin site, pathogenic microbes must replicate before causing full-blown infection and disease. Within cells, viral particles release their nucleic acids, which may be directly translated into viral proteins (positive-strand RNA viruses), transcribed from a negative strand of RNA into a complementary mRNA (negative-strand RNA viruses), or transcribed into a complementary strand of DNA (retroviruses); for DNA viruses, mRNA may be transcribed directly from viral DNA, either in the cell nucleus or in the cytoplasm. To grow, bacteria must acquire specific nutrients

or synthesize them from precursors in host tissues. For example, since aromatic amino acids are not available as nutrients in host tissues, pathogenic bacteria must synthesize them from precursors. Many infectious processes are usually confined to specific epithelial surfaces -- influenza to the respiratory mucosa, gonorrhea to the urogenital epithelium, shigellosis to the gastrointestinal epithelium. While there are multiple reasons for this specificity, one important consideration is probably the ability of these pathogens to obtain from these specific environments the nutrients needed for growth and survival.

Temperature restrictions also play a role in limiting certain pathogens to specific tissues. Rhinoviruses, a cause of the common cold, grow best at 33°C and replicate in cooler nasal tissues but not in the lung. Leprosy lesions due to *Mycobacterium leprae* are found in and on relatively cool body sites. Fungal pathogens that infect the skin, hair follicles, and nails (dermatophyte infections) remain confined to the cooler, exterior, keratinous layer of the epithelium.

A topic of major interest is the ability of many bacterial, fungal, and protozoal species to grow in multicellular masses referred to as *biofilms*. These masses are biochemically and morphologically quite distinct from the free-living individual cells referred to as *planktonic cells*. Growth in biofilms leads to altered microbial metabolism, production of extracellular virulence factors, and decreased susceptibility to biocides, antimicrobial agents, and host defense molecules and cells. Regulation of biofilm morphogenesis is controlled by bacterial quorum-sensing systems. Quorum sensing for some organisms involves the production of homoserine lactone molecules such as *N*-(3-oxododecanoyl)homoserine lactone and *N*-butyrylhomoserine lactone, which combine with transcriptional activators to control gene expression. *P. aeruginosa* growing on the bronchial mucosa during chronic infection, staphylococci and other pathogens growing on implanted medical devices, and dental pathogens growing on tooth surfaces to form plaques represent several examples of microbial biofilm growth associated with human disease. Many other pathogens can form biofilms during in vitro growth, including *Helicobacter pylori*, the cause of stomach ulcers; *E. coli* O157, one cause of the hemolytic-uremic syndrome; and *Gardnerella vaginalis*, an organism associated with bacterial vaginosis.

AVOIDANCE OF INNATE HOST DEFENSES

As microbes have probably interacted with mucosal/epithelial surfaces since the emergence of multicellular organisms, it is not surprising that multicellular hosts have a variety of innate surface defense mechanisms that can sense when pathogens are present and contribute to their elimination. The skin, a formidable physical barrier to microbial entry, is both acidic and bathed with fatty acids toxic to many microbes. Successful skin pathogens such as staphylococci must tolerate these adverse conditions. Mucosal surfaces themselves present a barrier composed of a thick mucus layer that entraps microbes and facilitates their transport out of the body by such processes as mucociliary clearance, coughing, and urination. Mucous secretions, saliva, and tears contain antibacterial factors such as lysozyme and antiviral factors such as interferons. Gastric acidity is inimical to the survival of many ingested pathogens, and many mucosal surfaces -- particularly the nasopharynx, the vaginal tract, and the gastrointestinal tract -- contain a resident flora of commensal microbes that interfere

with the ability of pathogens to colonize and infect a host.

Pathogens that survive these factors must still contend with host endocytic, phagocytic, and inflammatory responses as well as with host genetic factors that determine the degree to which a pathogen can survive and grow. The growth of viral pathogens entering skin or mucosal epithelial cells can be limited by a variety of host genetic factors, including production of interferons, modulation of receptors for viral entry, and age- and hormone-related susceptibility factors; by nutritional status; and even by personal habits such as smoking and exercise.

ENCOUNTERS WITH EPITHELIAL CELLS

Over the past decade, many bacterial pathogens have been shown to enter epithelial cells (Fig. 120-1), often using specialized surface structures that bind to receptors for internalization. However, the exact role and the importance of this process in infection and disease are not well defined for most of these pathogens. Bacterial entry into host epithelial cells is seen as a means for dissemination to adjacent or deeper tissues or as a route to sanctuary to avoid ingestion and killing by professional phagocytes. Epithelial cell entry appears, for instance, to be a critical aspect of dysentery induction by *Shigella*.

Curiously, the less virulent strains of many bacterial pathogens are more adept at entering epithelial cells than are more virulent strains; examples include pathogens that lack the surface polysaccharide capsule needed to cause serious disease. Thus, for *Haemophilus influenzae*, *Streptococcus pneumoniae*, *S. agalactiae* (group B *Streptococcus*), and *S. pyogenes*, isogenic mutants or variants lacking capsules enter epithelial cells better than the wild-type, encapsulated parental forms that cause disseminated disease. These observations have led to the proposal that epithelial cell entry may be a manifestation of host defense, resulting in bacterial clearance by both shedding of epithelial cells containing internalized bacteria and initiation of a subclinical inflammatory response. However, a consequence of this process would be the opening of a hole in the epithelium, potentially allowing uningested organisms to enter the submucosa. This scenario has been documented in murine *Salmonella typhimurium* infections and in experimental bladder infections with uropathogenic *E. coli*. In the latter system, bacterial pili mediate cell attachment to integral membrane glycoproteins called uroplakins that coat the host cells, resulting in exfoliation of the cells with attached bacteria. Subsequently, infection is produced by residual bacterial cells that invade the denuded epithelium. Perhaps at low bacterial inocula epithelial cell ingestion and subclinical inflammation are efficient means to eliminate pathogens, while at higher inocula a proportion of surviving bacterial cells enter the host tissue through the damaged mucosal surface and multiply, producing disease. Alternatively, failure of the appropriate epithelial cell response to a pathogen may allow the organism to survive on a mucosal surface where, if it avoids other host defenses, it can grow and cause a local infection. Along these lines, as noted above, *P. aeruginosa* is taken into epithelial cells byCFTR, a protein missing or nonfunctional in most severe cases of cystic fibrosis. The major clinical consequence of this disease is chronic airway-surface infection with *P. aeruginosa* in 80 to 90% of patients with cystic fibrosis. The failure of airway epithelial cells to ingest and promote the removal of *P. aeruginosa* has been proposed as a key component of the hypersusceptibility of these patients to chronic airway infection.

ENCOUNTERS WITH PHAGOCYTES

Phagocytosis of microbes is a major innate host defense that limits the growth and spread of pathogens. Phagocytes appear rapidly at sites of infection in conjunction with the initiation of inflammation. Ingestion of microbes by both tissue-fixed macrophages and migrating phagocytes probably accounts for the limited ability of most microbial agents to cause disease. A family of related molecules called *collectins*, *soluble defense collagens*, or *pattern recognition molecules* are found in blood (mannose-binding lectin), in lung (surfactant proteins A and D), and most likely in other tissues as well and bind to carbohydrates on microbial surfaces to promote phagocyte clearance. Bacterial pathogens seem to be ingested principally by polymorphonuclear neutrophils (PMNs), while eosinophils are frequently found at sites of infection by protozoan or multicellular parasites. Successful pathogens, by definition, must avoid being cleared by professional phagocytes. One of several antiphagocytic strategies employed by bacteria and by the fungal pathogen *Cryptococcus neoformans* is to elaborate large-molecular-weight surface polysaccharide antigens, often in the form of a capsule that coats the cell surface. Most pathogenic bacteria produce such antiphagocytic capsules.

As activation of local phagocytes in tissues is a key step in initiating inflammation and migration of additional phagocytes into infected sites, much attention has been paid to microbial factors that initiate inflammation. Encounters with phagocytes are governed largely by the structure of the microbial constituents that elicit inflammation, and detailed knowledge of these structures for bacterial pathogens has contributed greatly to our understanding of molecular mechanisms of microbial pathogenesis (Fig. 120-2). The best-studied system involves the interaction ofLPS from gram-negative bacteria and the glycosylphosphatidylinositol (GPI)-anchored membrane protein CD14 found on the surface of professional phagocytes, including migrating and tissue-fixed macrophages andPMNs. A soluble form of CD14 is also found in plasma and on mucosal surfaces. A plasma protein, LPS-binding protein (LBP), transfers LPS to membrane-bound CD14 on myeloid cells and promotes binding of LPS to soluble CD14. Soluble CD14/LPS/LBP complexes bind to many cell types and may be internalized to initiate cellular responses to microbial pathogens. It has been shown that peptidoglycan and lipoteichoic acid from gram-positive bacteria and cell-surface products of mycobacteria and spirochetes can interact with CD14 (Fig. 120-2).

GPI-anchored receptors do not have intracellular signaling domains, and mammalian Toll-like receptors (TLRs) transduce signals for cellular activation due toLPSbinding. TLRs initiate cellular activation through a series of signal-transducing molecules (Fig. 120-2) that lead to nuclear translocation of the transcription factor NF-kB, a master-switch for production of important inflammatory cytokines such as tumor necrosis factora (TNF-a) and interleukin (IL) 1.

The initiation of inflammation can occur not only with LPS and peptidoglycan but also with viral particles and other microbial products such as polysaccharides, enzymes, and toxins.

Bacterial Cell Wall Structure Gram-positive bacteria have a rigid cell wall that gives the organisms their characteristic shape, differentiates them from eukaryotic cells, and allows them to survive in osmotically unfavorable environments. The cell wall is

composed mainly of peptidoglycan (Fig. 120-3, panel *C*; a polymer of *N*-acetylglucosamine and its lactyl ether, *N*-acetylmuramic acid), with peptide side chains covalently bound to the lactyl group (Fig. 120-3, panel *A*). The peptide chains consist of alternating D and L amino acids and are usually linked to each other by a pentaglycine bridge binding a terminal D-alanine on one peptide substituent to the penultimate L-lysine on a neighboring peptide. Variations in this basic structure have been described for a number of bacterial genera. In addition, the cell walls of gram-positive bacteria contain teichoic acids (Fig. 120-3, panel *D*), phosphate-linked polymers of ribitol or glycerol that can have additional compounds linked to available side groups. Lipid tails anchor these acids to the cytoplasmic membrane, giving rise to lipoteichoic acids.

Gram-negative bacteria possess a cytoplasmic membrane and a peptidoglycan layer similar to but reduced from that found in gram-positive organisms, but these organisms also produce an outer membrane that is covalently linked to the tetrapeptides of the peptidoglycan layer by a lipoprotein (Fig. 120-3, panel *B*). Embedded in the outer membrane are special proteins with important functions, including maintaining the outer membrane's integrity, acting as a selective barrier for diffusion of molecules into the cell, serving as receptors for bacteriophages, and binding siderophores that scavenge iron for transport into the bacterial cell. The exterior layer of the outer membrane contains the major surface glycolipid, which can be either a classical bacterialLPS or a lipooligosaccharide (LOS); pathogens such as *Neisseria* and *Haemophilus* spp. express LOS, which contains smaller polysaccharide constituents. Although LPS/LOS was thought to be essential to the viability of gram-negative bacteria, a viable strain of *N. meningitidis* lacking LOS has now been made.

Exterior to theLPS for many, but not all, gram-negative pathogens is a capsular polysaccharide, which (along with LPS for some pathogenic species) confers resistance to phagocytosis by preventing innate host opsonins, such as the complement proteins C3 and C4, from coating the organisms -- a process that promotes their uptake by phagocytes. Capsular polysaccharides are also important extracellular components of gram-positive bacteria, serving as critical factors in bacterial resistance to opsonophagocytosis and phagocytic killing. Variation in the expression of the capsule in *S. pneumoniae* accounts for the different morphologies of colonies of this pathogen on agar plates (smooth and rough phenotypes); this property was exploited in studies proving that DNA carries genetic information in a cell.

Lipopolysaccharide Most of the important biologic properties associated withLPS(endotoxin) are due to the lipid A portion (Fig. 120-3, panel *E*), a relatively conserved, highly acylated di-*N*-acetylglucosamine backbone linkedb 1®6 and containing phosphate groups on the reducing 1 and nonreducing 4¢ carbons. Attached to carbon 6¢ is the inner polysaccharide core, which is usually, but not always, composed of a di- or trisaccharide of 2-keto-3-deoxyoctonate (KDO). Additional sugar substituents are linked to the inner core, forming a complete core. Attached to the complete core are either short polysaccharide side chains (formingLOS) or longer O polysaccharide side chains (forming complete LPS) composed of a variety of monosaccharides, substituted with a variety of components, such as formyl, acetyl, and hydroxy-butyryl side chains; amino acids or peptides; and phosphate groups. Further biologic functions of LOS/LPS that are important to the survival of microbes after entry

into a host include resistance to the bacteriolytic effects of complement and protection against antimicrobial factors such as defensins and bactericidal permeability-increasing protein, a molecule closely related toLBP in structure and function. Defensins are found in high concentrations in granules of myeloid cells, including platelets, and are usually highly cationic peptides capable of insertion into bacterial cells and killing of these cells.

Additional Interactions of Microbial Pathogens and Phagocytes Other ways that microbial pathogens avoid destruction by phagocytes include production of factors that are toxic to the phagocytes or that interfere with the chemotactic and ingestion function of phagocytes. Hemolysins, leukocidins, and the like are microbial proteins that can kill phagocytes that are attempting to ingest organisms elaborating these substances. For example, staphylococcal hemolysins inhibit macrophage chemotaxis and kill these phagocytes. Streptolysin O made by *S. pyogenes* binds to cholesterol in phagocyte membranes and initiates a process of internal degranulation, with the release of normally granule-sequestered toxic components into the phagocyte's cytoplasm. *Entamoeba histolytica*, an intestinal protozoan that causes amebic dysentery, can disrupt phagocyte membranes after direct contact via the release of protozoal phospholipase A and pore-forming peptides.

Microbial Survival Inside Phagocytes Many important microbial pathogens use a variety of strategies to survive inside phagocytes (particularly macrophages) after ingestion. Inhibition of fusion of the phagocytic vacuole (the phagosome) containing the ingested microbe with the lysosomal granules containing antimicrobial substances (the lysosome) allows *M. tuberculosis*, *S. typhi*, and *Toxoplasma gondii* to survive inside macrophages. Some organisms, such as *Listeria monocytogenes*, escape into the phagocyte's cytoplasm to grow and eventually spread to other cells. Resistance to killing within the macrophage and subsequent growth are critical to successful infection by herpes-type viruses, measles virus, poxviruses, *Salmonella*, *Yersinia*, *Legionella*, *Mycobacterium*, *Trypanosoma*, *Nocardia*, *Histoplasma*, *Toxoplasma*, and *Rickettsia*. *Salmonella* spp. use a master regulatory system, in which the *PhoP/PhoQ* genes control other genes, to enter and survive within cells, with intracellular survival entailing structural changes in the cell envelopeLPS.

TISSUE INVASION AND TISSUE TROPISM

TISSUE INVASION

Most viral pathogens cause disease by growth at skin or mucosal entry sites, but some pathogens spread from the initial site to deeper tissues. Virus can spread via the nerves (rabies virus) or plasma (picornaviruses) or within migratory blood cells (poliovirus, Epstein-Barr virus, and many others). Specific viral genes determine where and how individual viral strains can spread.

Bacteria may invade deeper layers of mucosal tissue via intracellular uptake by epithelial cells, traversal of epithelial cell junctions, or penetration through denuded epithelial surfaces. Among virulent *Shigella* strains and invasive *E. coli*, outer-membrane proteins are critical to epithelial cell invasion and bacterial multiplication. *Neisseria* and *Haemophilus* spp. penetrate mucosal cells by poorly understood mechanisms before dissemination into the bloodstream. Staphylococci and

streptococci elaborate a variety of extracellular enzymes, such as hyaluronidase, lipases, nucleases, and hemolysins, that are probably important in breaking down cellular and matrix structures and allowing the bacteria access to deeper tissues and blood. Organisms that colonize the gastrointestinal tract can often translocate through the mucosa into the blood and, under circumstances in which host defenses are inadequate, cause bacteremia. *Yersinia enterocolitica* can invade the mucosa through the activity of the invasin protein. Some bacteria (e.g., *Brucella*) can be carried from a mucosal site to a distant site by phagocytic cells (e.g.,PMNs) that ingest but fail to kill the bacteria.

Fungal pathogens almost always take advantage of host immunocompromise to spread hematogenously to deeper tissues. The AIDS epidemic has resoundingly illustrated this principle: the immunodeficiency of many HIV-infected patients permits the development of life-threatening fungal infections of the lung, blood, and brain. Other than the capsule of *C. neoformans*, specific fungal antigens involved in tissue invasion are not well characterized. Both fungal pathogens and protozoal pathogens (e.g., *Plasmodium* spp. and *E. histolytica*) undergo morphologic changes to spread within a host. Malarial parasites grow in liver cells as merozoites and are released into the blood to invade erythrocytes and become trophozoites. *E. histolytica* is found as both a cyst and a trophozoite in the intestinal lumen, through which this pathogen enters the host, but only the trophozoite form can spread systemically to cause amebic liver abscesses. Other protozoal pathogens, such as *T. gondii*, *Giardia lamblia*, and *Cryptosporidium*, also undergo extensive morphologic changes after initial infection to spread to other tissues.

TISSUE TROPISM

The propensity of certain microbes to cause disease by infecting specific tissues has been known since the early days of bacteriology, yet the molecular basis for this propensity is understood somewhat better for viral pathogens than for other agents of infectious disease. Specific receptor-ligand interactions clearly underlie the ability of certain viruses to enter cells within tissues and disrupt normal tissue function, but the mere presence of a receptor for a virus on a target tissue is not sufficient for tissue tropism. Factors in the cell, route of viral entry, viral capacity to penetrate into cells, viral genetic elements that regulate gene expression, and pathways of viral spread in a tissue all affect tissue tropism. Some viral genes are best transcribed in specific target cells, such as hepatitis B genes in liver cells and Epstein-Barr virus genes in B lymphocytes. The route of inoculation of poliovirus determines its neurotropism, although the molecular basis for this circumstance is not understood.

The lesser understanding of the tissue tropism of bacterial and parasitic infections is exemplified by *Neisseria* spp. There is no well-accepted explanation of why *N. gonorrhoeae* colonizes and infects the human genital tract while the closely related species *N. meningitidis* principally colonizes the human oropharynx. *N. meningitidis* expresses a capsular polysaccharide, while *N. gonorrhoeae* does not; however, there is no indication that this property plays a role in the different tissue tropisms displayed by these two bacterial species. *N. gonorrhoeae* can use cytidine monophosphate *N*-acetylneuraminic acid from host tissues to add *N*-acetylneuraminic acid (sialic acid) to itsLOS O side chain, and this alteration appears to make the organism resistant to host defenses. Lactate, present at high levels on genital mucosal surfaces, stimulates

sialylation of gonococcal LOS. Bacteria with sialic acid sugars in their capsules, such as *N. meningitidis*, *E. coli* K-1, and group B streptococci, have a propensity to cause meningitis, but this generalization has many exceptions. For example, all recognized serotypes of group B streptococci contain sialic acid in their capsules, but only one serotype (III) is responsible for most cases of group B streptococcal meningitis. Moreover, both *H. influenzae* and *S. pneumoniae* can readily cause meningitis, but these organisms do not have sialic acid in their capsules.

TISSUE DAMAGE AND DISEASE

Disease is a complex phenomenon resulting from tissue invasion and destruction, toxin elaboration, and host response. Viruses cause much of their damage by exerting a cytopathic effect on host cells and inhibiting host defenses. The growth of bacterial, fungal, and protozoal parasites in tissue, which may or may not be accompanied by toxin elaboration, can also compromise tissue function and lead to disease. For some bacterial and possibly some fungal pathogens, toxin production is one of the best-characterized molecular mechanisms of pathogenesis, while host factors such as IL-1,TNF-a, kinins, inflammatory proteins, products of complement activation, and mediators derived from arachidonic acid metabolites (leukotrienes) and cellular degranulation (histamines) readily contribute to the severity of disease.

VIRAL DISEASE

Viral pathogens are well known to inhibit host immune responses by a variety of mechanisms. Immune responses can be affected by down-regulating production of most major histocompatibility complex (MHC) molecules (adenovirus E3 protein), by diminishing cytotoxic T cell recognition of virus-infected cells (Epstein-Barr virus EBNA1 antigen and cytomegalovirus IE protein), by producing virus-encoded complement receptor proteins (herpesvirus and vaccinia virus) that protect infected cells from complement-mediated lysis, by making proteins that interfere with the action of interferon (influenza virus and poxvirus), and by elaborating superantigen-like proteins (mouse mammary tumor virus and related retroviruses, rabies nucleocapsid, and possibly the Nef protein of HIV). Superantigens activate large populations of T cells that express particular subsets of the T cell receptorb protein, causing massive cytokine release and subsequent host reactions. Another molecular mechanism of viral virulence involves the production of peptide growth factors for host cells, which disrupt normal cellular growth, proliferation, and differentiation. In addition, viral factors can bind to and interfere with the function of host receptors for signaling molecules. Modulation of cytokine production during viral infection can stimulate viral growth inside cells with receptors for the cytokine, and virus-encoded cytokine homologues (e.g., the Epstein-Barr virus BCRF1 protein, which is highly homologous to the immunoinhibitoryIL-10 molecule) can potentially prevent immune-mediated clearance of viral particles. Viruses can cause disease in neural cells by interfering with levels of neurotransmitters without necessarily destroying the cells, or they may induce either programmed cell death (apoptosis) to destroy tissues or inhibitors of apoptosis to allow for prolonged viral infection of cells. Overall, any disruption of normal cellular and tissue function due to viral infection can underlie the resultant clinical disease.

BACTERIAL TOXINS

Among the first infectious diseases to be understood were those due to toxin-elaborating bacteria. Diphtheria, botulism, and tetanus toxins are responsible for the diseases associated with local infections due to *Corynebacterium diphtheriae*, *Clostridium botulinum*, and *Clostridium tetani*, respectively. Enterotoxins produced by *E. coli*, *Salmonella*, *Shigella*, *Staphylococcus*, and *V. cholerae* contribute to diarrheal disease caused by these organisms. Staphylococci, streptococci, *P. aeruginosa*, and *Bordetella* elaborate various toxins that cause or contribute to disease, including toxic shock syndrome toxin 1 (TSST-1); erythrogenic toxin; exotoxins A, S, and U; and pertussis toxin. A number of these toxins (e.g., cholera toxin, diphtheria toxin, pertussis toxin, *E. coli* heat-labile toxin, and *P. aeruginosa* exotoxin) have adenosine diphosphate (ADP)-ribosyltransferase activity; i.e., the toxins enzymatically catalyze the transfer of the ADP-ribosyl portion of nicotinamide adenine diphosphate to target proteins and inactivate them. The staphylococcal enterotoxins, TSST-1, and the streptococcal pyogenic exotoxins behave as superantigens, stimulating certain T cells to proliferate without processing of the protein toxin by antigen-presenting cells. Part of this process involves stimulation of the antigen-presenting cells to produceIL-1 andTNF-a, which have been implicated in many of the clinical features of diseases like toxic shock syndrome and scarlet fever. A number of gram-negative pathogens (*Salmonella*, *Yersinia*, and *P. aeruginosa*) possess the ability to inject toxins directly into host target cells by means of a complex set of proteins referred to as the type III secretion system.

ENDOTOXIN

The lipid A portion of gram-negativeLPS has potent biologic activities that cause many of the clinical manifestations of gram-negative bacterial sepsis, including fever, muscle proteolysis, uncontrolled intravascular coagulation, and shock. The effects of lipid A appear to be mediated by the production of potent cytokines due to LPS binding to CD14 and signal transduction viaTLRs, particularly TLR4. Cytokines exhibit potent hypothermic activity through effects on the hypothalamus; they also increase vascular permeability, alter the activity of endothelial cells, and induce endothelial-cell procoagulant activity. Numerous therapeutic strategies aimed at neutralizing the effects of endotoxin are under investigation, but so far the results have been disappointing.

INVASION

Many diseases are caused primarily by pathogens growing in tissue sites that are normally sterile. Pneumococcal pneumonia is mostly attributable to the growth of *S. pneumoniae* in the lung and the attendant host inflammatory response, although specific factors that enhance this process (e.g., pneumolysin) may be responsible for some of the pathogenic potential of the pneumococcus. Disease that follows bacteremia and invasion of the meninges by meningitis-producing bacteria such as *N. meningitidis*, *H. influenzae*, *E. coli* K1, and group B streptococci appears to be due solely to the ability of these organisms to gain access to these tissues, multiply in them, and provoke cytokine production leading to tissue-damaging host inflammation.

Specific molecular mechanisms accounting for tissue invasion by fungal and protozoal pathogens are less well described. Except for studies pointing to factors like capsule and melanin production by *C. neoformans* and possibly levels of cell wall glucans in

some pathogenic fungi, the molecular basis for fungal invasiveness is not well defined. Melanism has been shown to protect the fungal cell against death caused by phagocyte factors such as nitric oxide, superoxide, and hypochlorite. Morphogenic variation and production of proteases (e.g., the *Candida* aspartyl proteinase) have been implicated in fungal invasion of host tissues.

If pathogens are effectively to invade host tissues (particularly the blood), they must avoid the major host defenses represented by complement and phagocytic cells. Bacteria most often avoid these defenses through their cell surface polysaccharides -- either capsular polysaccharides or long O-side-chain antigens characteristic of the smoothLPS of gram-negative bacteria. These molecules can prevent the activation and/or deposition of complement opsonins or limit the access of phagocytic cells with receptors for complement opsonins to these molecules when they are deposited on the bacterial surface below the capsular layer. Another potential mechanism of microbial virulence is the ability of some organisms to present the capsule as an apparent self antigen through molecular mimicry. For example, the polysialic acid capsule of group B *N. meningitidis* is chemically identical to an oligosaccharide found on human brain cells.

Immunochemical studies of capsular polysaccharides have led to an appreciation of the tremendous chemical diversity that can result from the linking of a few monosaccharides. For example, three hexoses can link up in more than 300 different, potentially serologically distinct ways, while three amino acids have only six possible peptide combinations. Capsular polysaccharides have been used as effective vaccines against meningococcal meningitis as well as against pneumococcal and *H. influenzae* infections and may prove to be of value as vaccines against any organisms that express a nontoxic, immunogenic capsular polysaccharide. In addition, most encapsulated pathogens become virtually avirulent when capsule production is interrupted by genetic manipulation; this observation emphasizes the importance of this structure in pathogenesis.

HOST RESPONSE

The inflammatory response of the host is critical for interruption and resolution of the infectious process but also is often responsible for the signs and symptoms of disease. Infection promotes a complex series of host responses involving the complement, kinin, and coagulation pathways. The production of cytokines such asIL-1,TNF-a, and other factors regulated in part by the NF-kB transcription factor leads to fever, muscle proteolysis, and other effects, as noted above. An inability to kill or contain the microbe usually results in further damage due to the progression of inflammation and infection. For example, in many chronic infections, degranulation of host inflammatory cells can lead to release of host proteases, elastases, histamines, and other toxic substances that can degrade host tissues. Chronic inflammation in any tissue can lead to the destruction of that tissue and to clinical disease associated with loss of organ function, such as sterility from pelvic inflammatory disease caused by chronic infection with *N. gonorrhoeae*.

The nature of the host response elicited by the pathogen often determines the pathology of a particular infection. Local inflammation produces local tissue damage, while systemic inflammation, such as that seen during sepsis, can result in the signs and

symptoms of septic shock. The severity of septic shock is associated with the degree of production of host effectors. Disease due to intracellular parasitism results from the formation of granulomas, wherein the host attempts to wall off the parasite inside a fibrotic lesion surrounded by fused epithelial cells that make up so-called multinucleated giant cells. A number of pathogens, particularly anaerobic bacteria, staphylococci, and streptococci, provoke the formation of an abscess, probably because of the presence of zwitterionic surface polysaccharides such as the capsular polysaccharide of *Bacteroides fragilis*. The outcome of an infection depends on the balance between an effective host response that eliminates a pathogen and an excessive inflammatory response that is associated with an inability to eliminate a pathogen and with the resultant tissue damage that leads to disease.

TRANSMISSION TO NEW HOSTS

As part of the pathogenic process, most microbes are shed from the host, often in a form infectious for susceptible individuals. However, the rate of transmissibility may not necessarily be high, even if the disease is severe in the infected individual, as these traits are not linked. Most pathogens exit via the same route by which they entered: respiratory pathogens by aerosols from sneezing or coughing or through salivary spread, gastrointestinal pathogens by fecal-oral spread, sexually transmitted diseases by venereal spread, and vector-borne organisms by either direct contact with the vector through a blood meal or indirect contact with organisms shed into environmental sources such as water. Microbial factors that specifically promote transmission are not well characterized. Respiratory shedding is facilitated by overproduction of mucous secretions, with consequently enhanced sneezing and coughing. Diarrheal toxins such as cholera toxin, *E. coli* heat-labile toxins, and *Shigella* toxins probably facilitate fecal-oral spread of microbial cells in the high volumes of diarrheal fluid produced during infection. The ability to produce phenotypic variants that resist hostile environmental factors (e.g., the highly resistant cysts of *E. histolytica* shed in feces) represents another mechanism of pathogenesis relevant to transmission. Blood parasites such as *Plasmodium* spp. change phenotype after ingestion by a mosquito -- a prerequisite for the continued transmission of this pathogen. Venereally transmitted pathogens may undergo phenotypic variation due to the production of specific factors to facilitate transmission, but shedding of these pathogens into the environment does not result in the formation of infectious foci.

In summary, the molecular mechanisms used by pathogens to colonize, invade, infect, and disrupt the host are numerous and diverse. Each phase of the infectious process involves a variety of microbial and host factors interacting in a manner that can result in disease. Recognition of the coordinated genetic regulation of virulence factor elaboration when organisms move from their natural environment into the mammalian host emphasizes the complex nature of the host-parasite interaction. Fortunately, the need for diverse factors in successful infection and disease implies that a variety of therapeutic strategies may be developed to interrupt this process and thereby prevent and treat microbial infections.

(Bibliography omitted in Palm version)

Back to Table of Contents

121. LABORATORY DIAGNOSIS OF INFECTIOUS DISEASES - *Andrew B. Onderdonk*

The laboratory diagnosis of infection requires the demonstration, either direct or indirect, of viral, bacterial, fungal, or parasitic agents in tissues, fluids, or excreta of the host. Clinical microbiology laboratories are responsible for processing these specimens and also for determining the antibiotic susceptibility of bacterial pathogens. Traditionally, detection of pathogenic agents has relied largely on either the microscopic visualization of pathogens in clinical material or the growth of microorganisms in the laboratory. Identification is generally based on phenotypic characteristics, such as fermentation profiles for bacteria, cytopathic effects in tissue culture for viral agents, and microscopic morphology for fungi and parasites. These techniques are reliable but are often time-consuming. Increasingly, the use of nucleic acid probes is becoming a standard detection and/or identification method in the clinical microbiology laboratory, gradually replacing phenotypic characterization and microscopic visualization methods.

DETECTION METHODS

Reappraisal of the methods employed in the clinical microbiology laboratory has led to the development of strategies for detection of pathogenic agents through nonvisual biologic signal detection systems. Much of this methodology is based on either computerization of detection systems with relatively inexpensive but sophisticated computers or the use of nucleic acid probes directed at specific DNA or RNA targets. This chapter discusses both the methods that are currently available and those that are being developed.

BIOLOGIC SIGNALS

A *biologic signal* is a material that can be reproducibly differentiated from other substances present in the same physical environment. Key issues in the use of a biologic (or electronic) signal are distinguishing it from background noise and translating it into meaningful information. Examples of biologic signals applicable to clinical microbiology include structural components of bacteria, fungi, and viruses; specific antigens; metabolic end products; unique DNA or RNA base sequences; enzymes; toxins or other proteins; and surface polysaccharides.

DETECTION SYSTEMS

A detector is used to sense a signal and to discriminate between the signal and background noise. Detection systems range from the trained eyes of a technologist assessing morphologic variations to sensitive electronic instruments, such as gas-liquid chromatographs coupled to computer systems for signal analysis. The sensitivity with which signals can be detected varies widely. It is essential to use a detection system that discerns small amounts of signal even when biologic background noise is present -- i.e., that is both sensitive and specific. Common detection systems include immunofluorescence; chemiluminescence for DNA/RNA probes; flame ionization detection of short- or long-chain fatty acids; and detection of substrate utilization or end-product formation as color changes, of enzyme activity as a change in light absorbance, of turbidity changes, of cytopathic effects in cell lines, and of particle

agglutination.

AMPLIFICATION

Amplification enhances the sensitivity with which weak signals can be detected. The most common microbiologic amplification technique is growth of a single bacterium into a discrete colony on an agar plate or into a suspension containing many identical organisms. The advantage of growth as an amplification method is that it requires only an appropriate growth medium; the disadvantage is the amount of time required for amplification. More rapid specific amplification of biologic signals can be achieved with techniques such as polymerase (ligase) chain reactions (PCRs, for DNA/RNA), enzyme immunoassays (EIAs, for antigens and antibodies), electronic amplification (for gas-liquid chromatography assays), antibody capture methods (for concentration and/or separation), and selective filtration or centrifugation. Although a variety of methods are available for the amplification and detection of biologic signals in research, thorough testing is required before they are validated as diagnostic assays.

DIRECT DETECTION

MICROSCOPY

The field of microbiology has been defined largely by the development and use of the microscope. The examination of specimens by microscopic methods rapidly provides useful diagnostic information. Staining techniques permit organisms to be seen more clearly.

The simplest method for microscopic evaluation is the wet mount, which is used, for example, to examine cerebrospinal fluid (CSF) for the presence of *Cryptococcus neoformans*, with India ink as a background against which to visualize large-capsuled yeast cells. Wet mounts with dark-field illumination are also used to detect spirochetes from genital lesions and to reveal *Borrelia* or *Leptospira* in blood. Skin scrapings and hair samples can be examined with use of either 10% KOH wet-mount preparations or the calcofluor white method and ultraviolet illumination to detect fungal elements as fluorescing structures. Staining of wet mounts -- for example, with lactophenol cotton blue stain for fungal elements -- is often used for morphologic identification. These techniques enhance signal detection and decrease the background, making it easier to identify specific fungal structures.

STAINING

Gram's Stain Without staining, bacteria are difficult to see at the magnifications (400 to 1000´) used for their detection. Although simple one-step stains can be used, differential stains are more common. Gram's stain differentiates between organisms with thick peptidoglycan cell walls (gram-positive) and those with outer membranes that can be dissolved with alcohol or acetone (gram-negative).

Gram's stain is particularly useful for examining sputum for polymorphonuclear leukocytes (PMNs) and bacteria. Sputum specimens with 25 or more PMNs and fewer than 10 epithelial cells per low-power field often provide clinically useful information.

However, the presence in "sputum" samples of more than 10 epithelial cells per low-power field and of multiple bacterial types suggests contamination with oral microflora. Despite the difficulty of discriminating between normal microflora and pathogens, Gram's stain may prove useful for specimens from areas with a large resident microflora if a useful biologic marker (signal) is available. Gram's staining of vaginal swab specimens can be used to detect epithelial cells covered with gram-positive bacteria in the absence of lactobacilli and the presence of gram-negative rods -- a scenario regarded as a sign of bacterial vaginosis. Similarly, examination of stained stool specimens for leukocytes is useful as a screening procedure before testing for *Clostridium difficile* toxin or other enteric pathogens.

The examination ofCSF and joint, pleural, or peritoneal fluid with Gram's stain is useful for determining whether bacteria and/orPMNs are present. The sensitivity is such that>10_4bacteria per milliliter should be detected. Centrifugation is often performed before staining to concentrate specimens thought to contain low numbers of organisms. The pellet is examined after staining. This simple method is particularly useful for examination of CSF for bacteria and white blood cells or of sputum for acid-fast bacilli (AFB).

Acid-Fast Stain The acid-fast stain identifies organisms that retain carbol fuchsin dye after acid/organic solvent disruption (e.g., *Mycobacterium* spp.). Modifications of this procedure allow the differentiation of *Actinomyces* from *Nocardia* or other weakly acid-fast organisms. The acid-fast stain is applied to sputum, other fluids, and tissue samples whenAFB(e.g., *Mycobacterium* spp.) are suspected. The identification of the pink/red AFB against the blue background of the counterstain requires a trained eye, since few AFB may be detected in an entire smear, even when the specimen has been concentrated by centrifugation. An alternative method is the auramine-rhodamine combination fluorescent dye technique.

Fluorochrome Stains Fluorochrome stains, such as acridine orange, are used to identify white blood cells, yeasts, and bacteria in body fluids. Other specialized stains, such as Dappe's stain, may be used for the detection of *Mycoplasma* in cell cultures. Capsular, flagellar, and spore stains are used for identification or demonstration of characteristic structures.

Immunofluorescent Stains The direct immunofluorescent antibody technique uses antibody coupled to a fluorescing compound, such as fluorescein, and directed at a specific antigenic target to visualize organisms or subcellular structures. When samples are examined under appropriate conditions, the fluorescing compound absorbs ultraviolet light and reemits light at a higher (visible) wavelength detectable by the human eye. In the indirect immunofluorescent antibody technique, an unlabeled (target) antibody binds a specific antigen. The specimen is then stained with fluorescein-labeled polyclonal antibody directed at the target antibody. Because each unlabeled target antibody attached to the appropriate antigen has multiple sites for attachment of the second antibody, the visual signal can be intensified (i.e., amplified). This form of staining is called *indirect* because a two-antibody system is used to generate the signal for detection of the antigen. Both direct and indirect methods detect viral inclusions (e.g., cytomegalovirus and herpes simplex virus) within cultured cells as well as many difficult-to-grow bacterial agents (e.g., *Legionella pneumophila*) directly in clinical

specimens.

MACROSCOPIC ANTIGEN DETECTION

Latex agglutination assays andEIAs are rapid and inexpensive methods for identifying organisms, extracellular toxins, and viral agents by means of protein and polysaccharide antigens. Such assays may be performed directly on clinical samples or after growth of organisms on agar plates or in viral cell cultures. The biologic signal in each case is the antigen to be detected. Monoclonal or polyclonal antibodies coupled to a reporter (such as latex particles or an enzyme) are used for detection of antibody-antigen binding reactions.

Techniques such as direct agglutination of bacterial cells with specific antibody are simple but relatively insensitive, while latex agglutination andEIAs are more sensitive. Some cell-associated antigens, such as capsular polysaccharides and lipopolysaccharides, can be detected by agglutination of a suspension of bacterial cells when antibody is added; this method is useful for typing of the somatic antigens of *Shigella* and *Salmonella*. In systems such as EIAs, which employ antibodies coupled to an enzyme, an antigen-antibody reaction results in the conversion of a colorless substrate to a colored product. Because the coupling of an enzyme to the antibody can amplify a weak biologic signal, the sensitivity of such assays is often high. In each instance, the basis for antigen detection is antigen-antibody binding, with the detection system changed to accommodate the biologic signal. Most such assays provide information as to whether antigen is present but do not quantify the antigen. EIAs are also useful for detecting bacterial toxins -- e.g., *C. difficile* toxins A and B in stool.

DETECTION OF PATHOGENIC AGENTS BY CULTURE

SPECIMEN COLLECTION AND TRANSPORT

To culture bacterial, mycotic, or viral pathogens, an appropriate sample must be placed into the proper medium for growth (amplification). The success of efforts to identify a specific pathogen often depends on the collection and transport process coupled to a laboratory-processing algorithm suitable for the specific sample/agent. In some instances, it is better for specimens to be plated at the time of collection rather than first being transported to the laboratory (e.g., urethral swabs being cultured for *Neisseria gonorrhoeae* or sputum specimens for pneumococci). In general, the more rapidly a specimen is plated onto appropriate media, the better the chance for isolating bacterial pathogens. Appendix B lists procedures for collection and transport of common specimens. Because there are many pathogen-specific paradigms for these procedures, it is important to seek advice from the microbiology laboratory when in doubt about a particular situation.

ISOLATION OF BACTERIAL PATHOGENS

Isolation of suspect pathogen(s) from clinical material relies on the use of artificial media that support bacterial growth in vitro. Such media are composed of agar, which is not metabolized by bacteria; nutrients to support the growth of the species of interest; and sometimes substances to inhibit the growth of other bacteria. Broth is employed for

growth (amplification) of organisms from specimens with few bacteria, such as peritoneal dialysis fluid,CSF, or samples in which anaerobes or other fastidious organisms may be present. The general use of liquid medium for all specimens is not worthwhile.

Two basic strategies are used to isolate pathogenic bacteria. The first is to employ enriched media that support the growth of any bacteria that may be present in a sample such as blood orCSF, which contain no bacteria under normal conditions. Broths that allow the growth of small numbers of organisms may be subcultured to solid media when growth is detected. The second strategy is to isolate (amplify) specific bacterial species from stool, genital tract secretions, or sputum -- sites that contain many bacteria under normal conditions. Antimicrobial agents or other inhibitory substances are incorporated into the agar medium to inhibit growth of all but the bacteria of interest. After incubation, organisms that grow on such media are further characterized to determine whether they are pathogens. Selection for organisms that may be pathogens from the normal microflora shortens the time required for diagnosis (Fig. 121-1).

ISOLATION OF VIRAL AGENTS (See also Chap. 180)

Pathogenic viral agents often are cultured when the presence of serum antibody is not a criterion for active infection or when an increase in serum antibody may not be detected during infection. The biologic signal -- virus -- is amplified to a detectable level. Although a number of techniques are available, an essential element is a monolayer of cultured mammalian cells sensitive to infection with the suspected virus. These cells serve as the amplification system by allowing the proliferation of viral particles. Virus may be detected by direct observation of the cultured cells for cytopathic effects or by immunofluorescent detection of viral antigens following incubation. Culture methods are particularly useful for detection of rapidly propagated agents, such as cytomegalovirus or herpes simplex virus.

AUTOMATION OF MICROBIAL DETECTION IN BLOOD

The detection of microbial pathogens in blood is difficult because the number of organisms present in the sample is often low and the organisms' integrity and ability to replicate may be damaged by humoral defense mechanisms or antimicrobial agents. Over the years, systems that rely on the detection of CO_2produced by bacteria and yeasts in blood culture medium have allowed the automation of the detection procedure. The most common systems involve either the insertion of a sampling device into each culture bottle at periodic intervals, with drawing off of the head-space gas for analysis by an infrared monitor, or the use of reflectance optics, with a light-emitting diode and photodiode employed to detect a color change in a CO_2-sensitive indicator built into the bottom of the culture bottle. These systems measure CO_2concentration as indicative of microbial growth. Sophisticated algorithms are used to evaluate the rate at which CO_2 is being produced and then to determine whether the rate of change is consistent with microbial growth. Such methods are no more sensitive than the human eye in detecting a positive culture; however, because the bottles in an automated system are monitored more frequently, a positive culture is often detected more rapidly than by manual techniques, and important information, including the result of Gram's stain and preliminary susceptibility assays, can be obtained sooner. One advantage of reflectance

optic systems is that the bottles are scanned continuously in a noninvasive monitoring procedure, and thus the likelihood of laboratory contamination is decreased.

Automated systems also have been applied to the detection of microbial growth from specimens other than blood, such as peritoneal and other normally sterile fluids. *Mycobacterium* spp. can be detected in certain automated systems if appropriate liquid media are used for culture.

DETECTION OF PATHOGENIC AGENTS BY SEROLOGIC METHODS

Measurement of serum antibody provides an indirect marker for past or current infection with a specific viral agent or other pathogens, including *Brucella*, *Legionella*, *Rickettsia*, and *Helicobacter pylori*. The biologic signal is usually either IgM or IgG antibody directed at surface-expressed antigen(s). The detection systems include those used for bacterial antigens (agglutination reactions, immunofluorescence, andEIA) and unique systems such as hemolysis inhibition and complement fixation. Serologic methods generally fall into two categories: those that determine protective antibody levels and those that measure changing antibody titers during infection. Determination of an antibody response as a measure of current immunity is important in the case of viral agents for which there are vaccines, such as rubella virus or varicella-zoster virus; assays for this purpose normally use one or two dilutions of serum for a qualitative determination of protective antibody levels. Quantitative serologic assays to detect increases in antibody titers most often employ paired serum samples obtained 10 to 14 days apart (i.e., acute- and convalescent-phase samples). Since the incubation period before symptoms are noted may be long enough for an antibody response to occur, the demonstration of acute-phase antibody alone is often insufficient to establish the diagnosis of active infection as opposed to past exposure. In such circumstances, IgM may be useful as a measure of an early, acute-phase antibody response. A fourfold increase in total antibody titer or in EIA activity between the acute- and convalescent-phase samples is also regarded as evidence for active infection.

For certain viral agents, such as Epstein-Barr virus, the antibodies produced may be directed at different antigens during different phases of the infection. For this reason, most laboratories test for antibody directed at both viral capsid antigens and antigens associated with recently infected host cells to determine the stage of infection.

IDENTIFICATION METHODS

Once bacteria are isolated, traditional methods of phenotypic characterization are often used to identify specific isolates. An organism's phenotypic characteristics include traits that are readily detectable after growth on agar media (colony size, color, hemolytic reactions, odor), use of specific substrates and carbon sources (such as carbohydrates), formation of specific end products during growth, and microscopic appearance. Broth tubes containing specific substrates are commonly employed for phenotypic characterization. While such methods have been used since the time of Pasteur, their simplicity and low cost continue to make them appealing today.

CLASSIC PHENOTYPING

Automated systems allow rapid phenotypic identification of bacterial pathogens. Most such systems are based on biotyping techniques, in which isolates are grown on multiple substrates and the reaction pattern is compared with known patterns for various bacterial species. This procedure is relatively fast, and commercially available systems include miniaturized fermentation, coding to simplify recording of results, and probability calculations for the most likely pathogens. If the biotyping approach is automated and the reading process is coupled to computer-based data analysis, rapidly growing organisms, such as Enterobacteriaceae, can be identified within hours of detection on agar plates.

Several systems use preformed enzymes for even speedier identification (within 2 to 3 h). Such systems do not rely on bacterial growth per se to determine whether a substrate has been used or not. They employ a heavy inoculum in which specific bacterial enzymes are present in sufficient quantity to convert substrate to product rapidly. In addition, some systems use fluorogenic substrate/end-product detection methods to increase sensitivity (through signal amplification).

GAS-LIQUID CHROMATOGRAPHY

Gas-liquid chromatography is often used to detect metabolic end products of bacterial fermentations. One common application is identification of short-chain fatty acids produced by obligate anaerobes during glucose fermentation. Because the types and relative concentrations of volatile acids differ among the various genera and species that make up this group of organisms, such information serves as a metabolic "fingerprint" for a particular isolate.

Gas-liquid chromatography can be coupled to a sophisticated signal-analysis software system for identification and quantitation of long-chain fatty acids (LCFAs) in the outer membranes and cell walls of bacteria and fungi. For any given species, the types and relative concentrations of LCFAs are distinctive enough to allow identification of even closely related species. An organism may be identified definitively within a few hours after detection of growth on appropriate media. LCFA analysis is one of the most advanced procedures currently available for phenotypic characterization.

NUCLEIC ACID PROBES

Techniques for the detection and quantitation of specific DNA and RNA base sequences in clinical specimens have become powerful tools for the diagnosis of bacterial, viral, parasitic, and fungal infections. The basic strategy is to detect a relatively short sequence of bases specific for a particular pathogen on single-stranded DNA or RNA by hybridization of a complementary sequence of bases (probe) coupled to a "reporter" system that serves as the signal for detection. Detection of an organism by nucleic acid probes offers a decided advantage over culture methods for difficult-to-grow organisms. Current technology encompasses a wide array of methods for amplification and signal detection, some of which have been approved by the U.S. Food and Drug Administration (FDA) for clinical diagnosis.

Use of nucleic acid probes generally involves lysis of intact cells and denaturation of the DNA or RNA to render it single-stranded. The probe may be hybridized to the target

sequence in a solution or on a solid support, depending on the system employed. In situ hybridization of a probe to a target is also possible and allows the use of probes with agents present in tissue specimens. Once the probe has been hybridized to the target (biologic signal), a variety of strategies may be employed to amplify and/or quantify the target-probe complex (Fig. 121-2).

Probes for Direct Detection of Pathogens in Clinical Specimens Nucleic acid probes are available commercially for direct detection of various bacterial and parasitic pathogens, including *L. pneumophila*, *Chlamydia trachomatis*, *N. gonorrhoeae*, group A *Streptococcus*, *Gardnerella vaginalis*, *Mycoplasma hominis*, and *Giardia lamblia*. In addition, probes for direct detection of human papillomavirus, *Candida* spp., and *Trichomonas vaginalis* have been approved. An assortment of probes for confirming the identity of cultured pathogens, such as *Mycobacterium* and *Salmonella* spp., are also available. Probes for the direct detection of bacterial pathogens are often aimed at highly conserved 16S ribosomal RNA sequences, of which there are many more copies than there are of any single genomic DNA sequence in a bacterial cell. The sensitivity and specificity of probe assays for direct detection are comparable to those of more traditional assays, includingEIA and culture. Many laboratories have developed their own probes for pathogens; however, unless a method-validation protocol for diagnostic testing has been performed, the use of such probes is restricted to research by federal law in the United States.

Nucleic Acid Probe Target-Amplification Strategies In theory, a single target nucleic acid sequence can be amplified to detectable levels. There are several strategies for target and/or probe amplification, includingPCR, ligase chain reaction, strand displacement amplification, and self-sustaining sequence replication. In each case, a target sequence or hybridized probe is amplified exponentially to obtain sufficient signal for detection, usually by the attachment of chemiluminescent reporter groups to the amplified product. The PCR strategy requires repeated heating of the DNA or RNA to separate the two complementary strands of the double helix, hybridization of a primer sequence to the appropriate target sequence, target amplification using the PCR for complementary strand extension, and signal detection via a labeled probe. The sensitivity of such assays is far greater than that of traditional assay methods such as culture. However, the care with which the assays are performed is important, because cross-contamination of clinical material with DNA or RNA from other sources (even at low levels) can cause false-positive results. An alternative method employs transcription-mediated amplification, in which an RNA target sequence is converted to DNA, which is then exponentially transcribed into RNA target. The advantage of this method is that only a single heating/annealing step is required for amplification. At present, amplification assays for *Mycobacterium tuberculosis*, *N. gonorrhoeae*, *C. trachomatis*, and *M. hominis* are on the market. Again, many laboratories have used commercially available *taq* polymerase, probe sequences, and reagents to develop "in-house" assays for diagnostic use. Issues related to quality control, interpretation of results, sample processing, and regulatory requirements have slowed the commercial development of diagnostic assay kits.

Signal Amplification Strategies Alternative systems for signal amplification have great appeal, particularly for quantitative determination of the amount of target present in a given specimen. With the advent of newer therapeutic regimens for HIV-associated

disease, cytomegalovirus infection, and hepatitis C virus infection, the response to therapy has been monitored by determining both genotype and "viral load" at various times after treatment initiation. Target amplification (PCR, transcription-mediated amplification) is difficult to control in a manner that allows accurate determination of the original target (genome) concentration. In other systems, probes attached to complementary target sequences are amplified by the attachment of a second probe and an amplification multimer to the original probe. In one such system, branched-chain DNA (bDNA)-based amplification, bDNA is attached to a site different from the target-binding sequence of the original probe. Chemiluminescent-labeled oligonucleotides can then bind to multiple repeating sequences on the bDNA. The amplified bDNA signal is detected by chemiluminescence. Alternatively, a DNA probe may be attached to an RNA target and the resulting DNA/RNA hybrid captured on a solid support by antibody specific for DNA/RNA hybrids (concentration/amplification) and detected by chemiluminescent-labeled antibody specific for DNA/RNA hybrids. Both methods can be used to determine the approximate number of target copies (virus) in the starting material. The advantage of these systems over PCR is that only a single heating/annealing step is required to hybridize the target-binding probe to the target sequence for amplification.

Application of Nucleic Acid Probe Technology Nucleic acid probe technology is being used to identify difficult-to-grow or noncultivable bacterial pathogens, such as *Mycobacterium*, *Legionella*, *Ehrlichia*, *Rickettsia*, *Babesia*, *Borrelia*, and *Tropheryma whippelii*. Amplification methods are also being used to detect chronic viral infections, such as herpes simplex encephalitis, cytomegalovirus infection, and hepatitis C. The monitoring of therapy with quantitative viral-load testing is a significant new application of nucleic acid technology. Further applications will likely include the replacement of culture for identification of many pathogens with solid-state DNA/RNA chip technology, in which thousands of unique nucleic acid sequences can be detected on a single computer chip. Probe technology also has the potential to detect viral pathogens faster than is possible with current culture techniques. However, if laboratories are to take full advantage of probe technology, the cost of reagents and assay automation must be competitive with the cost of existing methodology. At present, the detection of agents such as *C. trachomatis* or *N. gonorrhoeae* by probe technology is more expensive for most laboratories than detection by traditional culture orEIA. Moreover, because automated processing equipment is just beginning to find its way into the laboratory for these assays, nucleic acid amplification methods are both more labor-intensive and more expensive than other detection systems. In the absence of clear documentation of clinical utility, many laboratories continue to wait forFDAapproval of commercially available DNA/RNA probe assays rather than validating in-house assays.

SUSCEPTIBILITY TESTING

A principal responsibility of the clinical microbiology laboratory is to determine which antimicrobial agents inhibit a specific bacterial isolate. Such testing is used to screen for infection control problems, such as methicillin-resistant *Staphylococcus aureus*, vancomycin-resistant *Enterococcus faecium*, or extended-spectrum b-lactamase-producing organisms. Two approaches are useful. The first is a qualitative assessment of susceptibility, with responses categorized as susceptible, resistant, or intermediate. This approach can involve either the placement of paper disks containing

antibiotics on an agar surface inoculated with the bacterial strain to be tested (Kirby-Bauer or disk/agar diffusion method), with measurement of the zones of growth inhibition following incubation, or the use of broth tubes containing a set concentration of antibiotic (breakpoint method). These methods have been carefully calibrated against quantitative methods and clinical experience with each antibiotic, and zones of inhibition and breakpoints have been calculated on a species-by-species basis.

The second approach is to inoculate the test strain of bacteria into a series of broth tubes (or agar plates) with increasing concentrations of antibiotic. The lowest concentration of antibiotic that inhibits microbial growth in this test system is known as the *minimum inhibitory concentration* (MIC). If tubes in which no growth occurs are subcultured, the minimum concentration of antibiotic required to kill the starting inoculum can also be determined (*minimum bactericidal concentration*, or MBC). Quantitative susceptibility testing by the microbroth dilution technique, a miniaturized version of the broth dilution technique using microwell plates, lends itself to automation and is commonly used in larger clinical laboratories.

A novel version of the disk/agar diffusion method employs a quantitative diffusion gradient, or epsilometer (E-test), and uses an absorbent strip with a known gradient of antibiotic concentrations along its length. When the strip is placed on the surface of an agar plate seeded with a bacterial strain to be tested, antibiotic diffuses into the medium, and bacterial growth is inhibited. TheMIC is estimated as the lowest concentration that inhibits growth.

For some organisms, such as obligate anaerobes, routine susceptibility testing generally is not performed because of the difficulty of growing the organisms and the predictable sensitivity of most isolates to specific antibiotics.

(Bibliography omitted in Palm version)

Back to Table of Contents

122. IMMUNIZATION PRINCIPLES AND VACCINE USE - *Gerald T. Keusch*, *Kenneth J. Bart*

Most humans live their lives ignoring the certainty of their own mortality. Perhaps this fact explains why the adage "an ounce of prevention is worth a pound of cure" has so little effect on their everyday behavior. Even when it comes to acting to protect their young, parents are capable of ignoring the potential for mortality among their children (in the developed world) and of accepting the certainty of childhood deaths (in the developing world). In both settings, parents all too often fail to seek out and demand the best preventive measures available. Unless mandated by the law in the former setting or provided by benevolent organizations or governments in the latter, universal immunization has invariably remained an unattained goal. Compulsion and benevolence, it seems, are two essential components of immunization.

However, the integration of immunization practices (a major component of primary disease prevention) into routine health care services has provided caregivers with control over a substantial proportion of the disease and mortality that plagued the United States during the first half of the twentieth century (Table 122-1). For society today, immunization represents one of the most cost-effective means of preventing infectious disease. For every dollar spent, diphtheria/tetanus/pertussis (DTP) vaccine saves $29, measles/mumps/rubella (MMR) vaccine saves $21, trivalent oral poliovirus vaccine (OPV) saves $6, varicella vaccine saves $5, and *Haemophilus influenzae* type b vaccine saves $2. At present, >50 biologic products are licensed in the United States, and 6 vaccines (12 antigens) are used for routine immunization in the young, including diphtheria/tetanus/acellular pertussis vaccine (DTaP), inactivated poliovirus vaccine (IPV), MMR, *H. influenzae* type b (Hib) vaccine, hepatitis B virus (HBV) vaccine, and varicella vaccine. Five vaccines are designed for routine use in adults: tetanus/diphtheria (Td) toxoids, adsorbed, for adult use; HBV vaccine; influenza virus vaccine; polyvalent pneumococcal polysaccharide vaccine; and varicella vaccine. Some preparations are designated as special-use vaccines (e.g., hepatitis A vaccine for travelers). Unfortunately, vaccines for eukaryotic pathogens (protozoa and helminths), which affect a large proportion of the world's population, have been difficult to develop and remain only a hope for the future.

IMPACT OF IMMUNIZATION

The epidemiologically appropriate use of vaccines has resulted in the global eradication of smallpox and in the potential eradication of poliomyelitis in the next few years and of measles by 2020. Already achieved are the virtual elimination of congenital rubella syndrome, tetanus, and diphtheria as well as a dramatic reduction in pertussis, rubella, measles, and mumps in the United States. The introduction ofHibconjugate vaccines for immunization of infants has all but eliminated invasive *Haemophilus* infections (including meningitis and pneumonia), presumably because these vaccines also reduce nasopharyngeal carriage of Hib and induce protection before the period of greatest vulnerability in infancy. The recently licensed polyvalent pneumococcal polysaccharide conjugate vaccine promises to have the same impact on invasive pneumococcal disease, including otitis media.

DEFINITIONS

Vaccination and *immunization* are often used as interchangeable terms. However, the former denotes only the administration of a vaccine or toxoid, whereas the latter describes the process of inducing or providing immunity by any means, whether active or passive. Thus, vaccination does not guarantee immunization. *Active immunization* refers to the induction of immune defenses by the administration of antigens in appropriate forms, whereas *passive immunization* involves the provision of temporary protection by the administration of exogenously produced immune substances. Immunizing agents thus include vaccines, toxoids, and antibody-containing immunoglobulin preparations from human or animal donors (Table 122-2).

PRINCIPLES OF IMMUNIZATION

Artificial induction of immunity closely follows two well-tested principles of nature. The first, active immunization, can be traced at least as far back as Thucydides, who noted that people surviving epidemics of plague in Athens were spared during later outbreaks of the same disease. The second, passive immunization, is a natural process as well and is exemplified by the transplacental transmission of maternal antibodies to the fetus to provide protection against several diseases during the first months of life. Use of the two measures together may produce a complementary effect (as withHBVvaccine plus hepatitis B immune globulin) or may actually interfere with the development of immunity (as when measles vaccine is administered within 6 weeks of immunoglobulin). Depending on whether there are multiple species or serotypes of an organism and -- if so -- whether there are common, cross-reactive, protective antigens, a specific vaccine may induce protection against all representative forms of an infectious agent or against the immunizing strain only. One of the intrinsic virtues of whole-organism vaccines is that they potentially contain all protective antigens of the organism. However, this virtue is counterbalanced by an inherent problem with such vaccines: the possibility of adverse responses to reactive but nonprotective antigens present in the mix. Because the immune response to specific antigens is controlled genetically, all individuals cannot be expected to respond identically to the same vaccine.

APPROACHES TO ACTIVE IMMUNIZATION

The two standard approaches to active immunization are (1) the use of live, generally attenuated, infectious agents (e.g., measles virus); and (2) the use of inactivated agents or their constituents or products obtained by genetic recombination (e.g., acellular pertussis vaccines). For many diseases (e.g., poliomyelitis, influenza), both approaches have been employed. Live attenuated vaccines are believed to induce an immunologic response more nearly like that resulting from natural infection than the response induced by killed vaccines. Currently available inactivated or killed vaccines consist of inactivated whole organisms (e.g., plague vaccine); detoxified protein exotoxins (e.g., tetanus toxoid); recombinant protein antigens (e.g.,HBVvaccine); or carbohydrate antigens, either present as soluble purified capsular material (e.g., *Streptococcus pneumoniae* polysaccharides) or conjugated to a protein carrier (e.g.,Hibpolysaccharide conjugated to diphtheria or tetanus toxoids).

APPROACHES TO PASSIVE IMMUNIZATION

Passive immunization is generally used to provide temporary immunity in an unimmunized subject exposed to an infectious disease when active immunization either is unavailable (e.g., for cytomegalovirus infection) or has not been implemented before exposure (e.g., for rabies). Passive immunization is used in the treatment of certain disorders associated with toxins (e.g., diphtheria), in certain bites (those of snakes and spiders), and as a specific or nonspecific immunosuppressant [Rho(D) immune globulin and antilymphocyte globulin, respectively].

Three types of preparations are used in passive immunization: (1) standard human immune serum globulin for general use (e.g., gamma globulin), administered intramuscularly or intravenously; (2) special immune serum globulins with a known content of antibody for specific agents (e.g.,HBV or varicella-zoster immune globulin); and (3) animal sera and antitoxins.

ROUTE OF ADMINISTRATION

The route of administration in part determines the rapidity and nature of the immune responses to vaccines. Vaccines can be administered orally, intranasally, intradermally, subcutaneously, or intramuscularly. Parenterally administered vaccine may not induce mucosal secretory IgA, and mucosal immunization may not induce good systemic responses. Vaccines must be administered by the licensed route to ensure immunogenicity and safety. For example, administration ofHBVvaccine into the gluteal rather than the deltoid muscle often fails to induce an adequate immune response, while subcutaneous rather than intramuscular administration ofDTPincreases the risk of reactions.

AGE

Because age influences the response to vaccines, schedules for immunization are based on age-dependent responses determined empirically from clinical trials. The presence of high levels of maternal antibody and/or the immaturity of the immune system in the early months of life impairs the initial immune response to some vaccines (e.g., measles orHibpolysaccharide but notHBV). In the elderly, vaccine responses may be diminished because of natural waning of the immune system. Hence, larger amounts of an antigen may be required to produce the desired response (e.g., in vaccination against influenza).

ADJUVANT POTENTIATION

The immune response to some antigens is potentiated by the addition of adjuvants such as aluminum salts or, in the case of polysaccharides (e.g., the polyribose phosphate oligosaccharide of Hib), by conjugation to a carrier protein. Adjuvants, nonspecific boosters of immune responses, are used with inactivated products such as diphtheria and tetanus toxoids, acellular pertussis (aP) vaccine, andHBVvaccine. The mechanism for adjuvant enhancement of immunogenicity is not well defined but relates in part to the rendering of soluble antigens into a particulate form, the mobilization of phagocytes to the site of antigen deposition, and the slowing down of the release of antigens, which prolongs stimulation of the immune response.

THE IMMUNE RESPONSE

While many constituents of infectious microorganisms and their products, such as exotoxins, are or can be made to be immunogenic, only a limited number stimulate a protective immune response. The immune system is complex, and antigen composition and presentation are critical for stimulation of the desired immune responses.

The Primary Response In the primary response to a vaccine antigen, an apparent latent period of several days precedes the detection of humoral and cell-mediated immunity. Although the immune response is turned on by contact with the antigen and the immune system, measurable circulating antibodies do not appear for 7 to 10 days. The immunoglobulin class of the response also changes over time. Early-appearing IgM antibodies generally exhibit only low affinity for the antigen, whereas later-appearing IgG antibodies display high affinity. For "thymus-dependent" antigens, CD4+ T helper lymphocytes control the switch from IgM to IgG. Some individuals do not respond, even when presented repeatedly with a vaccine antigen, often because they lack the major histocompatibility complex determinants required to recognize the antigen. This situation is known as *primary vaccine failure*.

The Secondary Response Heightened humoral or cell-mediated responses are elicited by a second exposure to the same antigen. These secondary responses occur rapidly, usually within 4 or 5 days, and result, for example, in increased titers of IgG antibody. The secondary response depends on immunologic memory after the first exposure and is characterized by a marked proliferation of antibody-producing B lymphocytes and/or effector T cells. Polysaccharide vaccines, such as that for *S. pneumoniae*, evoke immune responses that are independent of T cells and are not enhanced by repeated administration. Covalent linking of polysaccharides to proteins converts the former to T cell-dependent antigens that induce immunologic memory and secondary responses to revaccination. Although levels of vaccine-induced antibodies may decline over time (*secondary vaccine failure*), revaccination or exposure to the organism may elicit a rapid protective secondary response consisting of IgG antibodies with little or no detectable IgM. This *anamnestic response* indicates that immunity has persisted. The lack of measurable antibody does not necessarily mean that the individual is unprotected. Furthermore, the mere presence of detectable antibodies after the administration of some vaccines and toxoids does not ensure clinical protection. A minimal circulating level of antibody is known to be required for protection from some diseases (e.g., 0.01 IU/mL for tetanus antitoxin).

Hypersensitivity Reactions Independent of antibody production, the stimulation of the immune system by vaccination may elicit unanticipated responses, especially hypersensitivity reactions. In the past, killed measles vaccine induced incomplete humoral immunity and cell-mediated hypersensitivity, resulting in the development of a syndrome of atypical measles in some children after subsequent exposure; thus this type of vaccine is no longer in use.

Mucosal Immunity Some pathogens are confined to and replicate only at mucosal surfaces (e.g., *Vibrio cholerae*), while others are able to penetrate the mucosa and replicate (e.g., poliovirus, rubella virus, and influenza virus). At the mucosal site, these organisms induce secretory IgA. The induction of secretory IgA by vaccines may be an

efficient way to block the essential first steps in pathogenesis, whether the organism is restricted to mucosal surfaces or systemically invades the host across mucosal surfaces.

Measurement of the Immune Response Immune responses to vaccines are often gauged by the concentration of specific antibody in serum. While seroconversion serves as a dependable indicator of an immune response, it measures only one immunologic parameter and does not necessarily indicate protection. The development of circulating antibodies after immunization often correlates directly with clinical protection (e.g., against measles or rubella). Some responses may not in themselves confer immunity but may be sufficiently associated with protection that they remain useful proxy measures of protective immunity (e.g., vibriocidal serum antibodies in cholera).

HERD IMMUNITY

It is not necessary to immunize every person in order to stop transmission of an infectious agent through a population. For those organisms dependent on person-to-person transmission, there may be a definable prevalence of immunity in the population above which it becomes difficult for the organism to circulate and reach new susceptibles. This prevalence is called *herd immunity*. When herd immunity is operative, the goals of immunization are converted from the immunization of every person in the community to the immunization of a specified minimum percentage of persons at risk. Herd immunity may wane if immunization capacity fails (as in diphtheria in the new independent states of the former Soviet Union) or if a sufficient percentage of individuals refuse to be immunized (as in pertussis in the United Kingdom and Japan in the 1970s because concern about infrequent -- albeit severe -- vaccine reactions came to exceed the fear of the disease itself). In both situations, loss of herd immunity led to renewed circulation of the organism and increased susceptibility to infection, with subsequent large outbreaks.

TARGET POPULATIONS AND TIMING OF IMMUNIZATION

For common and highly communicable childhood diseases like measles, the target population is the universe of susceptible individuals, and the time to immunize is as early in life as is feasible. Epidemiologic differences in measles in different settings, however, dictate different strategies of immunization. In the industrialized world, immunization with live-virus vaccine at 12 to 15 months of age has been the norm because the vaccine protects>95% of those immunized at this age and there is little measles morbidity/mortality among very young infants. In contrast, in the developing world, measles accounts for a significant proportion of deaths of young infants. Thus it is desirable to immunize children during the first few months of life in order to narrow the window of vulnerability between the rapid decline of maternal antibody after 4 to 6 months and the development of vaccine-induced active immunity.

Hibcauses meningitis, epiglottitis, and pneumonia in early childhood, with rates rising sharply after the disappearance of maternally derived antibody. The first Hib vaccines often failed when administered during infancy; this failure was due mainly to an age-related inability to respond to polysaccharide antigens. To overcome this problem, the protective polysaccharide was coupled to protein and converted to a T

cell-dependent antigen to which young infants could respond.

In contrast to measles andHibinfection, rubella is primarily a threat to the fetus; young infants and children are not at risk of serious illness. Given the susceptibility of the fetus, immunization of all women of reproductive age before pregnancy would be an ideal strategy. However, it is difficult to systematically vaccinate adolescent and young-adult females. Thus, to assure the protection of as many women as possible, the rubella component is included in a combination vaccine with mumps and measles (MMR) that is administered during infancy.

Some vaccines are now used primarily for adults. For example, influenza virus and polyvalent pneumococcal polysaccharide vaccines are used to prevent pneumonia deaths in the elderly. Unfortunately, these vaccines are underutilized, in part because physicians and otherwise healthy individuals in the target group ignore the indications and in part because there is still a tendency to think about disease prevention with vaccines as a strategy for children. Pneumococcal polysaccharide vaccine is also recommended for children>2 years old who are at risk of severe or even life-threatening pneumococcal infection, such as those with sickle cell disease, asplenia (whether functional or anatomic), renal failure with nephrotic syndrome, cerebrospinal fluid leak, and HIV infection or other immunosuppressive disease states.

THE DEVELOPMENT OF VACCINES

BIOLOGIC IMPEDIMENTS

There are often major technical problems to overcome in vaccine development. Although just one major antigenic type of influenza virus is typically in circulation at any one time, the virus is characterized biologically by its antigenic drift. Thus, a new antigenic version capable of causing a global pandemic emerges periodically, and a new vaccine must be rapidly devised, produced, and distributed. In contrast, many prevalent pneumococcal polysaccharide serotypes circulate at all times. Because immunity to the pneumococcus is serotype specific, an individual is susceptible to all serotypes against which he or she lacks antibody. Serotype-specific protection has made it more difficult to develop an effective pneumococcal vaccine than it was to develop a vaccine against *H. influenzae*, of which one capsular serotype (type b) is associated with nearly all cases of severe disease. To overcome this problem, pneumococcal vaccine currently includes 23 polysaccharides that represent ~80% of the virulent serotypes commonly encountered in the United States. Unfortunately, some serotypes are poorly immunogenic, and immunized individuals remain susceptible to the serotypes not included in the vaccine.

STRATEGY FOR VACCINE DEVELOPMENT

Vaccine development depends on the systematic application of a four-phase strategy: (1) studies in animals to identify protective antigen, (2) determination of how to present this antigen effectively to the immune system, (3) assessment of the safety and immunogenicity of the preparation in small and then in large human populations at various ages, and (4) evaluation of safety and efficacy in the target population. Each of these steps is simple in concept but difficult in execution, not least because of the

clinical trials necessary to assess safety and efficacy; failure at any level stops the process. Thus, in 1995,>190 candidate vaccines were under investigation, but just 5 new products were licensed in the United States. Progress in immunology has taught us much about the organization and function of the immune system (Chap. 305); it has also taught us that the immune system is complicated and that details of antigen composition and presentation are critical for stimulating desired immune responses.

Ultimately, vaccines for humans must be tested in humans. After initial animal studies and small phase 1 and 2 human studies to assess immune responses, optimal dosage, and safety, clinical trials of vaccine efficacy are performed, sometimes with informed volunteers who are challenged with a virulent strain. Larger clinical effectiveness trials in the community, typically involving 1000 to 10,000 vaccinees, may lead to application for licensure. Because of their limited size, however, these trials cannot be expected to detect rare adverse effects. Thus, licensing does not guarantee that a new vaccine is completely safe, and postlicensing monitoring is needed to ensure effectiveness and to document the occurrence of adverse events of low frequency. In 1999, the recently licensed rhesus rotavirus vaccine was withdrawn because postmarketing surveillance uncovered an association with a rare event in infants, intussusception of the bowel.

The development of vaccines goes beyond technology and proof of principle to issues such as development costs, manufacturers' liability and indemnity, perceived public health needs, and the likelihood that a product will be used or sold. Given the complex science required, the costs of vaccine development are high and success is uncertain, adding risk to the development decision. It is unfortunate that the one sure implication of uncertainty in vaccine development is increased cost. In addition, a rational assignment of costs for development between the public and private sectors in the United States has never been achieved.

VACCINE FORMULATIONS

Studies of clinical immunology have shown that living and dead antigens do not necessarily induce the same immune responses and that the requirements for the development of protective immunity differ with the organism. These insights, together with the refinement of epidemiologic concepts surrounding immunization, have changed the strategy of vaccine development. The goal is not only to select the correct antigens but also to ensure that the vaccines will result in the type of immune response needed for protection, whether the T cell-mediated activation of macrophages or the generation of cytotoxic T cells, B cell-mediated secretory IgA, or a particular IgG subtype response to a specific polysaccharide epitope.

Live vaccines consist of selected or genetically altered organisms that are avirulent or dramatically attenuated yet remain immunogenic. These agents are expected to cause a subclinical illness that mimics natural infection except for the lack of clinically significant disease. They offer the advantage of replication in vivo, which increases the antigenic load presented to the host's immune system; they may confer lifelong protection with one dose; they present all expressed antigens, thus overcoming immunogenetic restrictions in some hosts; they may reach the local sites most relevant to the induction of protective immunity; and they may produce important protective antigens in vivo that are not efficiently expressed in vitro.

Nonviable vaccines may fail to elicit mucosal IgA-mediated immunity, as they lack a delivery system that will effectively transport them to local antigen-processing cells. Moreover, except for pure polysaccharide antigens, these preparations must almost always be given in multiple doses to induce effective responses. However, killed vaccines can be extremely effective. For example, the nonviable hepatitis A vaccine formulation appears to be close to 100% effective in inducing protective immunity. Methods are under development to incorporate vaccine antigens into degradable polymers that may release antigen at predictable times after a single inoculation and simulate multiple injections over time of the same vaccine.

In spite of their advantages, live vaccines are not always to be preferred. For example, liveOPV is contraindicated for use in children with immune-deficiency diseases and in their adult contacts. In addition, even though killed poliovirus vaccine does not completely immunize the gut and can neither reduce the circulation of wild-type poliovirus nor immunize contacts of vaccine recipients, the United States has now switched to a four-dose schedule of this vaccine because of the risk of vaccine-associated polio posed by live OPV.

To create a deliverable vaccine, constituents other than the antigens are required (Table 122-3). These constituents can affect the immunogenicity, efficacy, and safety of a vaccine and can render one formulation superior to another.

NEW VACCINE APPROACHES

The first generation of vaccines included whole killed -- or, more recently, live -- attenuated microorganisms or partially purified microbial products, such as tetanus toxoid, that induced protective antibodies. The second generation of vaccines has taken advantage of molecular genetics and protein chemistry to isolate and manipulate purified proteins or components or subunits of organisms or to generate genetically engineered and attenuated live native organisms or cloned antigens expressed by harmless vector organisms. One conceptual leap is the production of transgenic plants expressing protective vaccine antigens (cloned, for example, in potatoes or bananas) that, when ingested orally, induce mucosal and systemic immune responses to homologous infectious challenges. While the practical use of this technique awaits further refinement, the concept that protective immunity can be induced in this manner has been proven in both animals and humans. Ease of production, stability, ease of administration without equipment, and low cost are the obvious advantages.

Another conceptual leap has led to a third generation of vaccines, in which nucleic acids (either DNA or RNA) are used to induce immunity. Development of DNA vaccines is at a more advanced stage. The principle is simple. First, a DNA plasmid containing the gene sequence for the immunogenic protein or fragment of interest is assembled, and the gene is placed under the control of a strong promoter and an appropriate transcription termination sequence. A single immunization with the plasmid (via intramuscular or intradermal injection, helium-accelerated gene gun injection of DNA-coated gold particles, compressed-air pneumatic jet injection of soluble DNA, direct skin application after suitable preparation, or even insertion of biodegradable stents loaded with the DNA of interest) results in DNA uptake into cells where the gene is expressed and

processed normally; thus the product stimulates an immune response. Alteration of the DNA construct or of the mode of administration or the coadministration of cytokine genes can determine whether the immune response is humoral or cellular or whether it involves primarily Th1, Th2, or cytotoxic T cells. Such decisions can be used to optimize the protective immunity induced.

This form of immunization offers real advantages and only theoretical and remote disadvantages (Table 122-4). Moreover, DNA vaccines may be useful in inducing tumor immunity, treating allergy (by suppressing IgE production), or even administering genes for gene therapy. RNA vaccines would avoid some of the potential concerns raised by DNA vaccines because RNA is less stable and does not persist or integrate into the chromosome or cause insertional mutagenesis. However, this lack of stability and the likely need for multiple doses, along with the increased cost of producing, storing, and transporting RNA, are significant disadvantages that remain to be overcome. The concept of nucleic acid vaccines -- whether based on DNA or RNA -- has been validated experimentally, and early human trials have begun. There is great optimism for the future but much to be learned if we are to apply this powerful new immunization technique successfully.

PRODUCTION OF VACCINES

As products to be given to healthy individuals to prevent disease, vaccines must not only be efficacious but also cause no harm. In the United States, quality assurance is the responsibility of vaccine manufacturers. Standards of manufacture of biologics [known as good manufacturing practices (GMPs)] are regulated and supervised by the U.S. Food and Drug Administration (FDA). Proof of the safety, efficacy, sterility, and purity of products is required before licensure, and sterility and purity are continually monitored for all lots of vaccine after licensure. Postmarketing studies of safety (phase IV studies) are part of routine regulatory control. On rare occasions, either GMP or quality assurance is inadequate; for example, the release of incompletely killed Salk polio vaccine in 1955 caused an outbreak of poliomyelitis in nearly 200 vaccine recipients and their contacts. Unregulated and uncontrolled manufacture of vaccines in developing countries has sometimes led to immunization with inactive products that fail to provide the expected protective immunity.

Another problem in the production of vaccines has unexpectedly cropped up in the past decade. For various reasons, including the high costs of vaccine development and the prospect of much higher profitability from investments in other products, the number of vaccine manufacturers in the United States has declined and the cost of some basic childhood vaccines has increased. Concern therefore exists about the future availability of these essential biologics for national use. Furthermore, pricing decisions made within the private-sector pharmaceutical industry can have a major impact on vaccine use. This situation has stimulated an initiative toward increased public involvement in supplying vaccine to individuals for whom price is an issue as well as in oversight of the vaccine supply and of price negotiations with the industry.

ADMINISTRATION OF VACCINES

Health care workers administering vaccines must take the precautions necessary to

minimize the risk of spreading disease -- for example, hand washing between immunizations. They should be immunized against hepatitis B, measles, rubella, influenza, and varicella. Different vaccines should not be mixed in the same syringe unless such a practice is specifically endorsed by licensure. Disposable needles and syringes should be discarded in labeled, puncture-proof containers to prevent inadvertent needlestick injury or reuse.

The addition of new, individually injectable vaccines to the immunization schedule has heightened parental concerns about the administration of up to four injections at a single clinic visit. The development and use of combinations of vaccines are intended to mitigate these concerns. Even when multiple injections are required, providers must make every effort to administer all indicated vaccines at each visit.

Wherever effective primary health care systems ensure access to medical services for the majority and the population is educated about the need for and efficacy of vaccines, coverage rates for basic immunization are usually high, regardless of the route of vaccine administration or the number of doses necessary. However, without systematic attention to the completion of multiple-dose vaccine schedules, coverage rates for second, third, and booster doses may drop off significantly.

USE OF VACCINES

Recommendations for vaccine use in the United States are developed by several different groups. These recommendations are the result of a collaborative process among the recommending groups, the pharmaceutical industry, and theFDA.

Vaccines recommended in 1999 for routine administration to infants, children, and adults are shown in Table 122-5; vaccines recommended for special use are shown in Table 122-6; and schedules for immunization of children and adults are shown in Fig. 122-1 andTable 122-7, respectively. The recommendations on route, site, and dosages for vaccination are derived from theoretical considerations, experimental trials, and clinical experience; deviation from these recommendations can result in inadequate protection. The administration of doses at intervals longer than those recommended does not diminish the ultimate protective response but merely delays it. It is not necessary to restart an interrupted series from the beginning or to add an extra dose. In contrast, giving vaccines at shorter-than-recommended intervals may result in poor responses.

RECORDING AND REPORTING REQUIREMENTS

Certain aspects of vaccine use are regulated by the National Childhood Vaccine Injury Act (NCVIA) of 1986 (modified in 1995). The act requires that all mandated childhood vaccinations be recorded by health care providers in the child's permanent medical record, including date of administration, manufacturer and lot number, and name of the provider administering the vaccine. State-based immunization information systems and registries are being developed to help public and private providers manage their immunization activities and particularly to address the problem of assessing immunization coverage when an individual's records are divided among multiple medical facilities.

Parents must be informed about the benefits and risks of immunization and should maintain an up-to-date immunization record on their children. Educational materials providing the required information (Vaccine Information Statements, VISs) are available from the American Academy of Pediatrics (AAP) or the Centers for Disease Control and Prevention (CDC).

VACCINES FOR ROUTINE USE

Infants and Children Recommended routine-use vaccines and schedules for their administration to infants and children are shown in Table 122-5 and Fig. 122-1, respectively. It is current practice for all children in the United States to receive DTaP, poliovirus, MMR, Hib, HBV, and varicella vaccines unless there are specific contraindications. Hepatitis A vaccine is currently recommended when there is a special risk of exposure to infection due to residence in communities with elevated rates of hepatitis A or travel to highly endemic countries.

Adults (See Table 122-7) All adults should be immune to diphtheria and tetanus. If not previously immunized, adults require a primary immunizing course of Td. Many individuals remain immune to tetanus into adulthood because they have received tetanus toxoid rather than Td after injuries, but they are commonly at risk of diphtheria because of the decline in titer of diphtheria antitoxin and the lack of boosting against diphtheria. The development of acellular pertussis vaccines that appear to be safe in adults may lead to a recommendation for booster immunization of adults if clinical trials confirm safety and efficacy. Routine immunization against polio is not recommended for adults unless they are at particular risk of exposure (e.g., through travel to endemic regions, as discussed below) or are the parents or guardians of a child with an immunodeficiency disorder. Adults should be protected from measles, mumps, and rubella; they should be vaccinated unless they are known to have received vaccine on or after their first birthday or to have had physician-diagnosed disease. Rubella vaccine should be given to all women of childbearing age unless they have documentary proof of immunization after their first birthday or laboratory evidence of immunity. An unsupported history of rubella disease is unreliable and should not be accepted. Adults without a clear history of chickenpox should receive varicella vaccine. College students, particularly freshmen living in a dormitory, are at increased risk of meningococcal meningitis. They should be made aware of the polysaccharide vaccine for serogroups A, C, Y, and W-135 and should be offered the option of immunization.

Current recommendations also include influenza vaccine for routine annual administration to adults³55 years of age and to individuals with chronic illness at any age. Polyvalent pneumococcal polysaccharide vaccine is similarly recommended for the elderly or chronically ill. HBV vaccine is recommended for individuals at high risk of exposure, including health care workers exposed to potentially infected blood or blood products, homosexuals, injection drug users, individuals living and working in institutions for the mentally retarded, and household contacts of known carriers of hepatitis B surface antigen (HBsAg). A new recombinant outer-surface protein A (rOspA) is licensed for persons 15 to 70 years of age for Lyme disease (LYMErix, SmithKline Beecham Pharmaceuticals), with use based on individual risk (geography and risk of exposure to ticks).

Adverse Events Modern vaccines, while safe and effective, are associated with adverse effects that range from infrequent and very mild to rare and life-threatening. The decision to use a vaccine involves an assessment of the risks of disease, the benefits of vaccination, and the risks associated with vaccination. Because these factors may change over time, continued assessment is essential.Table 122-8 lists valid and invalid contraindications to immunization and describes appropriate precautions in the use of specific vaccines. Antivaccine advocacy groups actively encourage avoidance of immunization because of their unproven belief that vaccines may cause certain disorders (for example, autism). This situation presents a challenge to the physician in educating parents about vaccine benefits and risks.

Vaccine components, including protective antigens, animal proteins introduced during vaccine production, and antibiotics or other preservatives or stabilizers, can cause allergic reactions in some recipients. These reactions may be local or systemic and may include urticaria and serious anaphylaxis. The most common extraneous allergen is egg protein introduced when vaccines such as those for measles, mumps, influenza, and yellow fever are prepared in embryonated eggs. Local or systemic reactions can result from too-frequent administration of vaccines such asTd, diphtheria/tetanus (DT), or rabies; these reactions are probably due to antigen-antibody complexes. In addition, live-virus vaccines can interfere with tuberculin test responses. When a tuberculin skin test is indicated, it should be done either on the day of immunization or 6 weeks later. When influenza vaccine is given to children <13 years old, only "split-virus" preparations should be used since whole-virus vaccines are associated with higher rates of adverse reactions in young children. Cumulative exposure to mercury in thimerosal-preserved vaccines is a concern, and plans are under way to replace current vaccines with thimerosal-free products. In the interim, infants born to HBsAg-negative mothers should not receive the initial dose of HBV vaccine at birth.

All detected adverse events temporally related to vaccination are expected to be reported to both the local health department and the vaccine manufacturer. TheNCVIArequires health care providers to report certain suspected adverse events following the receipt of a mandated vaccine to theFDA's Vaccine Adverse Events Reporting System (Table 122-9). Although a temporal relationship does not establish cause and effect, this surveillance system remains the only mechanism for collecting the data needed to form conclusions and make decisions.

USE OF VACCINES IN SPECIAL CIRCUMSTANCES

Pregnancy Because of theoretical risk to the fetus and real risk of litigation to the practitioner, routine immunization of pregnant women is best avoided. However, wherever hygienic conditions during delivery cannot be guaranteed, it is essential to ensure that pregnant women are immune to tetanus: the transfer of maternal antitoxin is an important means of preventing neonatal tetanus, and pregnant women can safely receive tetanus as well as diphtheria toxoids. Although live-virus vaccines in general should be withheld during pregnancy, polio and yellow fever vaccines are exceptions and may be administered if the risk of exposure to disease is great. If indicated, some inactivated vaccines (e.g.,HBV, influenza, and pneumococcal vaccines) are safe for pregnant women. Known pregnancy is considered a contraindication to the receipt of

rubella, measles, mumps, and varicella vaccines. Although of theoretical concern, no cases of congenital rubella syndrome or abnormalities attributable to rubella vaccine virus have been observed in infants born to susceptible mothers who received rubella vaccine during pregnancy.

Breast Feeding Neither killed nor live vaccine affects the safety of breast feeding for either mother or infant. Breast-fed infants can be immunized on a normal schedule.

Occupational Exposure Immunization recommendations for most occupational groups remain to be developed. Specific practices are now mandated by the Occupational Safety and Health Administration for the immunization of health care workers against hepatitis B in the United States. Rubella is transmitted to and from health care workers in medical facilities, particularly in pediatric practice. Health care workers who might transmit rubella to pregnant patients should be immune to rubella; it is prudent to screen these employees for antibodies to rubella virus and to immunize susceptible individuals. Persons providing health care are also at greater risk from measles and varicella than the general public, and those who are likely to come into contact with measles- and varicella-infected patients should be immune. Persons employed in caring for patients with chronic diseases can transmit influenza; such workers should be vaccinated annually. Unfortunately, these recommendations often are not fully implemented, even in academic institutions.

HIV Infection and Other Immunocompromised States Limited studies in HIV-infected individuals have found no increase in the risk of adverse events from live or inactivated vaccines. However, immune responses may not be as vigorous in immunocompromised individuals as in those with a normal immune system. Persons known to be infected with HIV should be immunized with recommended vaccines in the same manner as individuals with a normal immune system and as early in the course of their disease as possible, before immune function becomes significantly impaired. Live attenuatedMMRvaccine can be administered to this group, butOPVcannot (Table 122-10).IPVshould be used when vaccination against polio is indicated. Household contacts of immunocompromised individuals should be immune to polio; when vaccinated, they should receive IPV. In practice, it is not necessary to test for HIV before making decisions about the immunization of asymptomatic individuals from known HIV risk groups.

Live attenuated vaccines are normally contraindicated in immunocompromised patients, including those with congenital immunodeficiency syndromes and those receiving immunosuppressive therapy. Passive immunization with immunoglobulin preparations or antitoxins can be considered in individual cases, either as postexposure prophylaxis or as part of the treatment of established infection.

Postexposure Immunization For certain infections, active or passive immunization soon after exposure prevents or attenuates disease expression. Recommended postexposure immunization regimens are compiled in Table 122-11. Measles immune globulin given within 6 days of exposure may prevent or modify infection, and measles vaccine given within the first few days after exposure may prevent symptomatic infection. Although clinical manifestations of rubella in pregnant women are minimized by postexposure passive immunization, this approach may not prevent maternal

viremia, fetal infection, and congenital rubella syndrome. Therefore, the administration of immune globulin is recommended only for women developing rubella during pregnancy who will not consider abortion under any circumstances. Tetanus immune globulin can be used in patients with tetanus. Survivors with no history of tetanus immunization should receive a primary series of toxoid injections since disease does not result in the development of protective levels of antitoxin. Administration of rabies immune globulin plus rabies vaccine in the immediate postexposure period is highly effective in preventing disease. Similarly, for persons who have not been actively immunized, the use of immune globulin within 2 weeks of exposure to hepatitis A is likely to prevent clinical illness. Good data indicate the efficacy of human hepatitis B immune globulin in preventing disease after exposure. While no high-titer preparation is available for postexposure protection against non-A, non-B hepatitis, standard human immune serum globulin is efficacious.

Simultaneous Administration of Multiple Vaccines The simultaneous administration of the most widely used live and inactivated vaccines has not resulted in impaired antibody responses or in increased rates of adverse reactions. Simultaneous administration of vaccines is advantageous in that it increases the probability that a child will ultimately be fully immunized; it is also useful in any age group when the potential exists for exposure to multiple infectious diseases during travel to endemic countries. However, combinationDTaP/Hibvaccines should not be used for primary immunization of infants because the response to Hib is blunted and suboptimal. Live-virus vaccines not given together on the same day should generally be administered at least 30 days apart.

High doses of immune globulin may inhibit the efficacy of measles and rubella vaccines, and an interval of at least 3 months is recommended between the administration of immune globulin and that ofMMRvaccine or its components. Postpartum vaccination of rubella-susceptible women should not be delayed because of the administration of anti-Rho(D) immune globulin or any other blood product during the last trimester or at delivery. Should administration of an immune globulin preparation become necessary after vaccination, it should be postponed, if possible, for at least 14 days to allow time for vaccine-virus replication and development of immunity. In general, there is little interaction of immune globulin with inactivated vaccines, and postexposure passive prophylaxis can be given together withHBVvaccine or tetanus toxoid, resulting in both immediate and long-lasting protection.

Travel The International Sanitary Regulations allow countries to impose requirements for yellow fever and killed cholera vaccines as a condition for admission, even though the latter is not an effective public health tool. Travelers should know whether these vaccines are required for entry into the countries on their itinerary to avoid being turned back or immunized on the spot. Infants, children, and adults should have all routine immunizations updated before traveling, with particular attention to polio, measles, andDTP/DTaPor Tdvaccines. The use of hepatitis A vaccine may be advisable for travelers to some locales. Special-use vaccines (Table 122-6), including rabies, meningococcal polysaccharide, typhoid (oral live or Vi polysaccharide), Japanese encephalitis, and plague vaccines, should be considered for those individuals who expect to go beyond the usual tourist routes or to spend extended periods in rural areas in disease-endemic regions. Most U.S. cities have at least one travel clinic that

maintains up-to-date epidemiologic information and can provide the appropriate vaccines.

DELIVERY OF VACCINES

Over the past 25 years, considerable progress has been made to ensure that every child in the United States is fully immunized by the time of school entry. All 50 states now require immunization for school entry, and most have laws addressing attendance at preschools and day-care centers. The impact of immunization and of other improvements in the health care provided to the American population on the incidence of vaccine-preventable illness is shown inTable 122-1. Nonetheless, many children are not fully immunized, especially in poor and underserved communities. The failure to vaccinate preschool children was largely responsible for the resurgence of measles between 1989 and 1991, with >55,000 cases and >130 measles-related deaths. Outbreaks of pertussis, mumps, and congenital rubella syndrome have occurred for the same reason: low immunization rates among preschool children.

ACCESS TO IMMUNIZATION

Four major barriers to infant and childhood immunization have been identified within the health care system: (1) low public awareness and lack of public demand for immunization, (2) inadequate access to immunization services, (3) missed opportunities to administer vaccines, and (4) inadequate resources for public health and preventive programs. These problems are sources of public concern, and their solution is a priority for national health policy in the United States. In response, the Children's Immunization Initiative was begun in 1990. At the national level, this program includes outreach and educational campaigns to promote parental awareness of the value of vaccination and to encourage health care providers to use every opportunity to vaccinate the children in their care. At the state and local levels, community and business groups, religious and service groups, schools, and the media have joined together in community-based networks. A National Immunization Week each April has been established to focus attention on the vaccination needs of infants and children. To improve the quality and quantity of vaccination services, expanded immunization-clinic hours and computerization of immunization records are being implemented as well.

There has been only modest progress towards the goals for adult immunization in the United States. Adult-immunization goals are important: As many as 60,000 adults die each year of vaccine-preventable diseases for which effective vaccines are not being optimally used. Most persons ³65 years of age do not receive influenza vaccine each year, and even fewer have ever received pneumococcal vaccine. Health care providers more often miss vaccination opportunities with adults than with infants and children. From 60 to 90% of adults hospitalized for or dying of influenza-associated respiratory disease have received medical care during the previous year and could have been immunized at that time. Medicare reimbursement for excess hospitalization during influenza epidemics ranges from $750 million to $1 billion. Additional efforts are required to ensure that adults receive pneumococcal,Td, andHBVvaccines as well.

A special setting for adult immunization is the administration of certain vaccines to pregnant women to enhance passive immunity in their offspring (e.g., tetanus toxoid). In

most cases the mother herself derives important benefits as well. Immunization of the mother should be undertaken at least 6 weeks before delivery to allow for efficient transplacental transfer of antibody to the fetus.

STANDARDS FOR IMMUNIZATION PRACTICES

National standards of immunization for adult and pediatric practice have been established to define common policies and practices for public health clinics and in physicians' private offices (Table 122-12). These guidelines highlight the need to distinguish between valid contraindications and conditions that are often considered but are not in fact contraindications (Table 122-8). Among the valid contraindications applicable to all vaccines are a history of anaphylaxis or other serious allergic reactions to a vaccine or vaccine component and the presence of a moderate or severe illness, with or without fever. Infants who develop encephalopathy within 72 h of a dose ofDTP orDTaPshould not receive further doses; those who develop a "precaution" (Table 122-8) should not normally receive further doses. Because of theoretical risks to the fetus, pregnant women should not receiveMMR or varicella vaccine. Diarrhea, minor respiratory illness with or without fever, mild to moderate local reactions to a previous dose of vaccine, the concurrent or recent use of antimicrobial agents, mild to moderate malnutrition, or the convalescent phase of an acute illness are not valid contraindications to routine immunization. Failure to vaccinate children because of these conditions is increasingly viewed as a missed opportunity for immunization.

BIOTERRORISM

The end of the twentieth century witnessed a rise in the risk of bioterrorism. While smallpox has been eradicated, known stocks of smallpox virus still exist in the United States and Russia, and unknown stocks probably exist in other countries considered likely to engage in terrorism. Global supplies of smallpox vaccine for use in case of deliberate release of smallpox virus are inadequate, and millions of people are likely to become infected and die in the event of such a release. Steps are only now being taken to ensure sufficient stockpiles of vaccine for this eventuality, and it will be several more years before these reach a critical size.

THE NATIONAL VACCINE INJURY COMPENSATION PROGRAM

The use of mandated vaccines benefits society as a whole by reducing morbidity and the cost of care for preventable diseases and by reducing childhood mortality. Thus, in the United States, society has assumed the obligation to care for those injured by the administration of mandated vaccines. TheNCVIA of 1986 (modified in 1995) is the instrument in use to ensure fairness to injured persons as well as protection for federal, state, and local immunization programs; private immunization providers; and vaccine manufacturers. The act was designed to implement two vital public policies: (1) to provide prompt and fair compensation to the families of children who have died or have been injured as a result of routine mandated immunization; and (2) to reduce the adverse impact of the tort system on vaccine supply, cost, and innovation/development. The success of immunization programs in the United States depends upon the continued viability of the National Vaccine Injury Compensation Program.

CONTROL OF VACCINE-PREVENTABLE DISEASE

A continuing task of public health practice is to maintain individual and herd immunity. The job is not over once a population is fully vaccinated; rather, it is imperative to immunize each subsequent generation as long as the threat of the disease persists. Ongoing surveillance and prompt reporting of disease to local or state health departments are essential to this goal, ensuring a continuing awareness of the possibility of vaccine-preventable illness. Nearly all vaccine-preventable diseases are now notifiable, and individual case data are routinely forwarded to theCDC. These data are used to detect outbreaks or other unusual events that require investigation and to evaluate prevention and control policies, practices, and strategies.

As a direct consequence of successes in immunization, vaccine-preventable diseases have become less visible; ironically, this situation may foster complacency among parents and health care providers about routine immunization of children. Even among the affluent and educated, immunization levels may be low, reflecting a misunderstanding of the continuing threat of disease with which parents and health care providers have limited experience or perhaps an unjustifiably greater fear of adverse reactions to vaccination than of the potential for illness and death due to vaccine-preventable diseases. Health care workers play an essential role in influencing the attitudes of patients regarding appropriate immunization; therefore, it is essential that these professionals continually update their own knowledge about vaccines and about the epidemiology of vaccine-preventable illnesses.

RESEARCH ON VACCINES AND IMMUNIZATION

The potential to eradicate selected diseases and to build sustainable immunization programs that reach every child is not being fulfilled with existing vaccines and delivery technology. New vaccines or new formulations that will not only improve protective responses but also simplify the immunization schedule are needed. The ideal would be vaccines that can be administered orally early in life, that provide lifelong protection against multiple infections, that can be given as one or only a few doses, and that are less reactive and more heat stable than current vaccines. To attain these ambitious goals may take decades, but progress is already being made in the development of new combinations of current vaccines to facilitate complete immunization. The results will be applicable to immunization programs in both developed and developing countries.

REEMERGENCE OF CONTROLLED DISEASE AND EMERGENCE OF NEW DISEASE

The emergence of new pathogens is fostered by the genetic potential of microbes to evolve as well as by rapid changes in human demographics and behavior and in a global ecology that creates new or more favorable hosts. Proof of the need for continuing vaccine research is found in the emergence of new infectious diseases such as HIV infection, Lyme borreliosis, hantavirus pulmonary syndrome, and hepatitis C; the appearance of a new epidemic cholera strain (serotype O139 Bengal) that exhibits no cross-immunity with the traditional O1 serotype; and the increase in global incidence and in drug resistance of familiar diseases that were once considered under control, such as tuberculosis and malaria. In addition, some common illnesses without a

previously known etiology, such as peptic ulcer disease and cervical and nasopharyngeal cancer, have now been epidemiologically linked to specific infectious agents and have thus become potentially vaccine-preventable conditions.

DEVELOPMENT OF NEW VACCINES

For many serious or even life-threatening infectious diseases, no effective vaccines are available. Although many new vaccines are undergoing human trials, the task of developing vaccines is proving very complex. Priorities for the United States currently include research on the following vaccines: HIV, pneumococcus (conjugate), group B *Streptococcus*, respiratory syncytial virus, rotavirus, *Mycobacterium tuberculosis*, herpes simplex virus, influenza A and B viruses, and hepatitis C virus. Also of high priority are vaccines for two virus-associated tumors: cervical cancer (human papillomavirus) and nasopharyngeal cancer (Epstein-Barr virus).

INTERNATIONAL CONSIDERATIONS

Since the establishment of the World Health Organization's Expanded Programme on Immunization in 1981, levels of coverage for the recommended basic children's vaccines (bacille Calmette-Guerin, polio,DTP/DTaP, measles, andHBV) have risen from 5% to ~80% worldwide. Each year, at least 2.7 million deaths from measles, neonatal tetanus, and pertussis and 200,000 cases of paralysis due to polio are prevented by immunization. Despite the successes of this program, many vaccine-preventable diseases remain prevalent in the developing world. Measles, for example, continues to kill an estimated 1.5 million children each year, and cases of diphtheria, whooping cough, polio, and neonatal tetanus still occur at unacceptably high rates. It is estimated that between 20 and 35% of all deaths of children under the age of 5 years are still associated with vaccine-preventable diseases.

In addition to the antigens included in the Expanded Programme for routine use in the developing world, others (Hib, Japanese B encephalitis, yellow fever, group A meningococcus, mumps, and rubella) are used regionally, depending on disease epidemiology and resources. Polio has been targeted for eradication; this disease has already been eliminated in the Americas and Europe and is close to elimination in the Western Pacific.

Because infectious diseases know no geographic or political boundaries, uncontrolled disease anywhere in the world poses a threat to the United States. Vaccines offer the opportunity to control and even eradicate some diseases, and eradication means that vaccines are no longer needed. Vaccines represent the best hope for stopping the pandemic of HIV infection throughout the world. The experience with smallpox has shown that the eradication of disease is a remarkably good economic investment. The entire sum that the United States spent for the global smallpox eradication campaign has been recouped, in 1968 dollars, every 2.5 months since 1971. The global eradication of polio will save the United States over $300 million a year in vaccine and associated delivery costs and will save over $1.5 billion a year worldwide.

Issues of cost, liability, risk, and profitability limit the interest of the pharmaceutical industry in the development of vaccines (e.g., for malaria) that will be used primarily in

poor developing countries. Efforts have been made to create partnerships in public research and privately funded development. Activities of established international organizations (such as the World Health Organization) and some new organizations (such as the Global Alliance for Vaccines and Immunization, the International AIDS Vaccine Initiative, and the Bill and Melinda Gates Foundation) have helped to move the process forward with strategy development and implementation or new funding. New international schemes are being considered by wealthy industrial nations; for example, advance-purchase schemes are being proposed in which the purchase of effective vaccines is guaranteed, ensuring the profitability that the marketplace has provided for industry in the wealthy countries. The effectiveness of such approaches remains to be seen, but they offer much-needed hope for at-risk populations around the world.

SOURCES OF INFORMATION ON IMMUNIZATION

· Official vaccine package circulars and Vaccine Administration Statements from the Centers for Disease Control and Prevention

· Report of the Committee on Infectious Diseases of the American Academy of Pediatrics ("Red Book")

· Recommendations of the Advisory Committee on Immunization Practices, Centers for Disease Control and Prevention

· Guide for Adult Immunization, American College of Physicians

· Health Information for International Travel (published yearly) and Advisory Memoranda on Travel (published periodically), Centers for Disease Control and Prevention

· Control of Communicable Diseases in Man, American Public Health Association

· Technical Bulletin of the College of Obstetrics and Gynecology

· National Network for Immunization Information, Infectious Diseases Society of America/Pediatric Infectious Diseases Society/American Academy of Pediatrics/American Nurses Association

(Bibliography omitted in Palm version)

Back to Table of Contents

123. HEALTH ADVICE FOR INTERNATIONAL TRAVEL - *J.S. Keystone*, *P.E. Kozarsky*

In 1991, the World Health Organization estimated that more than 30 million persons traveled from industrialized countries to the developing world. Studies show that between 50 and 75% of short-term travelers to the tropics or subtropics report some health impairment. Most of these health problems are minor, with only 5% requiring medical attention and fewer than 1% requiring hospitalization.

Although infectious agents contribute substantially to morbidity in travelers, these pathogens account for only about 1% of deaths among this population. Cardiovascular disease and injuries are the most frequent causes of death among travelers from the United States, accounting for 49 and 22% of deaths, respectively. Age-specific rates of mortality due to cardiovascular disease are similar to those among nontravelers. In contrast, rates of death due to injury -- the majority from motor vehicle, drowning, or aircraft accidents -- are several times higher among travelers. Figure 123-1 summarizes the monthly incidence of health problems during travel in developing countries.

GENERAL ADVICE

Staying healthy during travel requires familiarity with the health risks that may be encountered at a given destination. However, health maintenance recommendations are based not only on the traveler's destination but also on risk assessment, which is determined by health status, specific itinerary, and lifestyle during travel. Detailed information regarding country-specific risks and recommendations may be obtained from many sources, including those listed under "Sources of Information on Travel Medicine."

IMMUNIZATIONS FOR TRAVEL

Immunizations for travel are generally divided into three categories: routine (childhood/adult boosters that are necessary regardless of travel), required (immunizations that are mandated by international regulations for entry into certain areas or for border crossings), and recommended (immunizations that are desirable because they confer protection against a variety of illnesses for which travel increases the risk). Vaccines commonly given to travelers are listed inTable 123-1.

Routine Immunizations

Diphtheria, Tetanus, and Polio Diphtheria continues to be a problem worldwide, with large outbreaks in the independent states formerly encompassed by the Soviet Union. Serosurveys show that tetanus antitoxin is lacking in many North Americans, especially those over the age of 50. Although the risk of polio to the international traveler is extremely low and although polio has been eradicated from the western hemisphere, studies in the United States have found varying levels of immunity in the general population; data indicate that 12% of adult American travelers are unprotected against at least one poliovirus serogroup. Foreign travel offers an ideal opportunity to have these immunizations updated.

Measles Measles (rubeola) continues to be a major cause of morbidity and mortality in the developing world (Chap. 194). Several outbreaks of measles in the United States have been linked to imported cases. The group at highest risk consists of persons born after 1956 and vaccinated before 1980, in many of whom primary vaccination failed. Travelers in this group should be reimmunized.

Influenza Influenza occurs year-round in the tropics and during the summer months in the southern hemisphere (which coincide with the winter months in the northern hemisphere). Vaccination should be considered for all travelers to these regions, particularly those who are elderly or chronically ill. The largest outbreak of travel-related influenza occurred in the summer of 1998 in Alaska and the Northwest Territories of Canada among cruise ship passengers and staff (Chap. 190).

Pneumococcal Infection Pneumococcal vaccine should be administered routinely to persons at high risk of serious pneumococcal infection, such as individuals with chronic heart, lung, or renal disease and those who have been splenectomized or have sickle cell disease.

Required Immunizations

Yellow Fever Documentation of yellow fever vaccination may be required for entry into countries of sub-Saharan Africa and equatorial South America, where the disease is endemic or epidemic, or into countries that are at risk of having the infection introduced. This vaccine is given only by state-authorized yellow fever centers, and its administration must be documented on an official International Certificate of Vaccination. The incidence of yellow fever among travelers is extremely low, probably because the vaccine is highly efficacious.

Cholera According to the World Health Organization, cholera vaccine should no longer be required for entry into any country. (However, see *recommended* use below.)

Meningococcal Meningitis Meningococcal vaccination is required for entry into Saudi Arabia during the Hajj. (For information on *recommended* use, see below.)

Recommended Immunizations

Hepatitis A and B Hepatitis A is the most frequent vaccine-preventable infection of travelers; the incidence of symptomatic infection during a 1-month stay in a developing country ranges from 3 to 6 cases per 1000. The risk is six times greater for those who stray from the usual tourist routes. The mortality rate for hepatitis A increases with age, reaching almost 3% among symptomatic individuals over age 50. Several vaccines are available in North America, each of which has an efficacy rate of >95%. The monthly incidence of hepatitis B infection, both symptomatic and asymptomatic, is 80 to 240 cases per 100,000. For reasons that are not entirely clear, long-stay overseas workers are at considerable risk for hepatitis B infection. In the near future, a combined hepatitis A and B vaccine should become available in the United States.

Typhoid Fever The attack rate for typhoid fever is 1 case per 30,000 per month of travel to the developing world (Chap. 156). However, the rates in India, Senegal, and North

Africa are tenfold higher, and, within these areas, rates are especially high among travelers to relatively remote destinations and among persons who are returning to their homelands to stay with relatives or friends. Each of the three available vaccines -- one oral and two injectable -- has an efficacy rate of approximately 70%.

Meningococcal Meningitis Although the risk of meningococcal disease among travelers has not been quantified, it is likely to be higher among those who live with poor indigenous populations in overcrowded conditions. The vaccine is recommended for persons traveling to sub-Saharan Africa during the dry season or to areas of the world where there are epidemics. Meningococcal vaccine, which protects against serogroups A/C/Y/W-135, has an efficacy rate of>90%.

Japanese Encephalitis The risk of Japanese encephalitis, an infection transmitted by mosquitoes in rural Asia and Southeast Asia, is approximately 1 case per 5000 per month of stay in an endemic area. Most symptomatic infections in U.S. residents have involved military personnel or their families. The vaccine efficacy rate is >80%. Serious allergic reactions sometimes occur; these reactions may be delayed in onset, developing up to 10 days after immunization. The vaccine is recommended for persons staying>1 month in endemic areas.

Cholera The risk of cholera is extremely low, with approximately 1 case per 500,000 journeys to endemic areas. Cholera vaccine is rarely recommended but should be considered for aid workers in refugee camps or in disaster/war-torn areas. The injectable vaccine available in the United States is only 30 to 50% effective, and its protective effect persists for only a short period. A more effective oral cholera vaccine is available in other countries.

Rabies Many cases of rabies have been reported in travelers, but there are no data on the risk of infection. Domestic animals are the major transmitters of rabies in developing countries (Chap. 197). Countries where canine rabies is highly endemic include Mexico, the Philippines, Sri Lanka, India, Thailand, and Vietnam. Each of the three vaccines available in the United States provides>90% protection. Rabies vaccine is recommended for long-stay travelers, particularly children, and persons who may be occupationally exposed in endemic areas.

PREVENTION OF MALARIA AND OTHER INSECT-BORNE DISEASES

It is estimated that more than 30,000 American and European travelers develop malaria each year (Chap. 214). Nevertheless, several studies indicate that fewer than 50% of U.S. travelers to malaria-endemic regions adhere to basic recommendations for malaria prevention.

The risk of malaria is highest in sub-Saharan Africa and Oceania (1:50 to 1:1000) and during the past decade has increased by more than fivefold for travelers to Kenya. The risk is intermediate (1:1000 to 1:12,000) for travelers to Haiti and the Indian subcontinent and is low (<1:50,000) for travelers to Asia and to Central and South America. Of the 1000 cases of malaria reported annually in the United States, 90% of those due to *Plasmodium falciparum* occur in travelers returning or immigrating from Africa and Oceania. With the worldwide increase in chloroquine- and multidrug-resistant

falciparum malaria, decisions about chemoprophylaxis have become more difficult. Moreover, the spread of malaria due to primaquine- and chloroquine-resistant strains of *P. vivax* has added to the complexity of treatment. The case-fatality rate of falciparum malaria in the United States is 4%; however, in only one-third of patients who die is the diagnosis of malaria considered before death.

Compliance with chemoprophylaxis regimens and use of personal protection measures are keys to the prevention of malaria. Personal protection measures are aimed at preventing mosquito bites, especially between dusk and dawn, and include the use of DEET-containing insect repellents, permethrin-impregnated bed nets and clothing, screened sleeping accommodations, and protective clothing. These practices also help prevent other insect-transmitted illnesses, such as dengue fever. Over the past decade, the incidence of dengue has increased, particularly in the Caribbean region, Latin America, and Southeast Asia. Since dengue fever is transmitted by a day-biting, urban-dwelling mosquito, attention to personal protection measures is required around the clock in most tropical areas.

The decision about whether or not to use malaria prophylaxis is based on the traveler's destination; the particular medication prescribed is determined not only by destination but also by the traveler's preference and medical history.Table 123-2 lists the currently recommended drugs of choice for prophylaxis of malaria by destination. Alternative medications for prophylaxis that are in use by some physicians include primaquine, atovaquone/proguanil, and chloroquine/proguanil.

PREVENTION OF GASTROINTESTINAL ILLNESS

Diarrhea is the leading cause of illness in travelers (Chap. 131) and is usually a short-lived, self-limited condition; however, 40% of affected individuals must alter their scheduled activities, and another 20% are confined to bed. The most important determinant of risk is the traveler's destination. Incidence rates per 2-week stay have been reported to be as low as 8% in industrialized countries and as high as 55% in parts of Africa, Central and South America, and Southeast Asia. Infants and young adults are at particularly high risk. The incidence of diarrhea is proportional to the number of dietary indiscretions. Studies of U.S. students in Mexico showed that eating meals in restaurants and cafeterias or consuming food from street vendors was associated with increased risk.

The most frequently identified pathogen causing traveler's diarrhea is toxigenic *Escherichia coli*, although in some parts of the world (notably northern Africa and Southeast Asia) *Campylobacter* infections appear to predominate. Other common causative organisms include *Salmonella*, *Shigella*, rotavirus, and the Norwalk agent. Except for giardiasis, parasitic infections are uncommon causes of traveler's diarrhea. A growing problem for travelers is the development of antibiotic resistance in many bacterial pathogens, including strains of *Campylobacter* and *Salmonella* resistant to quinolones and strains of *E. coli*, *Shigella*, and *Salmonella* resistant to trimethoprim-sulfamethoxazole.

Although the mainstay of prevention of traveler's diarrhea involves food and water precautions, the literature has repeatedly documented dietary indiscretions in 98% of

travelers within the first 48 h after arrival at their destination. The old maxim "Boil it, cook it, peel it, or forget it!" is easy to remember but appears to be difficult to adhere to. General food and water precautions include eating foods piping hot; avoiding foods that are raw, poorly cooked, or sold by street vendors; and drinking boiled or commercially bottled beverages, particularly those that are carbonated. Heating kills diarrhea-causing organisms, whereas freezing does not; therefore, ice cubes made from unpurified water should be avoided.

As traveler's diarrhea can occur despite rigorous food and water precautions, travelers should carry medications for self-treatment. For mild to moderate diarrhea, loperamide and fluid replacement may be sufficient. An antibiotic is useful in reducing the frequency of bowel movements and duration of illness in moderate to severe diarrhea. The standard regimen is a 3-day course of a quinolone taken twice daily or (in the case of some of the newer agents) once daily. However, studies have shown that a single large dose of a quinolone may be as effective as the 3-day regimen, particularly if infection with a multidrug-resistant organism is not suspected. For diarrhea acquired in areas such as Thailand, where>70% of *Campylobacter* infections are quinolone resistant, azithromycin may be a better choice.

Prophylaxis of traveler's diarrhea with bismuth subsalicylate is widely used but only about 60% effective. For certain individuals (e.g., athletes who must be in peak physical condition to perform, persons with a repeated history of traveler's diarrhea, and some persons with chronic diseases), a single daily dose of a quinolone antibiotic during travel of <1 month's duration is highly effective.

PREVENTION OF OTHER TRAVEL-RELATED PROBLEMS

Travelers are at high risk for *sexually transmitted diseases*. Surveys have shown that large numbers engage in casual sex, and there is a reluctance to use condoms consistently. An increasing number of travelers are being diagnosed with *schistosomiasis*. Travelers should be cautioned to avoid bathing, swimming, or wading in freshwater lakes, streams, or rivers in the parts of tropical South America, the Caribbean, Africa, and Southeast Asia where this infection can be acquired. Travelers are cautioned to avoid walking barefoot because of the risk of *hookworm* and *strongyloidiasis* and, at night, *snakebites*. Prevention of *travel-associated injury* depends mostly on common-sense precautions. Riding on motorcycles and overcrowded public vehicles is not recommended; in particular, individuals should not travel by road after dark in rural areas. In addition to its association with motor vehicle accidents, excessive alcohol use has been a significant factor in drownings, assaults, and injuries.

THE TRAVELER'S MEDICAL KIT

A traveler's medical kit is strongly advisable, particularly for long-stay travelers. The contents may vary widely, depending on the itinerary, duration of stay, style of travel, and local medical facilities. While many medications are available abroad, often over the counter, directions for their use may be nonexistent or in a foreign language, and, more important, a product may be outdated or counterfeit. Therefore, if possible, a complete supply of medications should accompany the traveler. In the kit, the short-term traveler should consider carrying an analgesic, an antidiarrheal agent, antihistamines, a laxative,

oral rehydration salts, sunscreen with a skin protection factor of at least 30, insect repellents (DEET) for the skin, an insecticide for clothing (permethrin), and (if necessary) an antimalarial. To these medications the long-stay traveler might add a broad-spectrum general-purpose antibiotic, an antibacterial eye and skin ointment, and a topical antifungal cream. Regardless of the duration of travel, a first-aid kit containing items such as scissors, tweezers, and bandages should be considered.

TRAVEL AND SPECIAL HOSTS

PREGNANCY AND TRAVEL

A woman's medical history and itinerary, the quality of medical care at her destinations, and her degree of flexibility determine whether travel is wise during pregnancy. According to the American College of Obstetrics and Gynecology, the safest part of pregnancy in which to travel is the second trimester (between 18 and 24 weeks), when there is the least danger of spontaneous abortion or premature labor. Some obstetricians prefer that women stay within a few hundred miles of home after the 28th week of pregnancy in case problems arise; in general, however, healthy women may be advised that it is acceptable to travel.

Despite this general recommendation, there are some relative contraindications to international travel during pregnancy, including certain obstetric risk factors: a history of miscarriage, premature labor, incompetent cervix, or toxemia. General medical problems such as diabetes, heart failure, severe anemia, or a history of thromboembolic disease should also prompt the pregnant woman to postpone her travels. Finally, regions in which the pregnant woman and her fetus may be at excessive risk (e.g., those at high altitudes and those where live-virus vaccines are required or where multidrug-resistant malaria is endemic) are not ideal destinations during any trimester.

Malaria Malaria during pregnancy carries a significant risk of morbidity and death. Levels of parasitemia are highest and failure to clear the parasites after chloroquine treatment are most frequent among primigravidae. Severe disease, with complications such as cerebral malaria, massive hemolysis, and renal failure, is especially likely in pregnancy. Fetal sequelae include spontaneous abortion, stillbirth, preterm delivery, and congenital infection.

Traveler's Diarrhea Because dehydration due to traveler's diarrhea can lead to inadequate placental blood flow, pregnant travelers must be extremely cautious regarding their food and beverage intake. The exclusive consumption of bottled (carbonated) or boiled drinks without ice, the eating of well-cooked meats and pasteurized dairy products, and the avoidance of pre-prepared salad items should help protect against traveler's diarrhea due to the usual causes as well as against infections such as toxoplasmosis, hepatitis E, and listeriosis, which can have serious sequelae in pregnancy.

The mainstay of therapy for traveler's diarrhea is rehydration. Kaolin-pectin combinations and loperamide may be used if necessary, but many of the usual antibiotics (e.g., quinolones) are contraindicated during pregnancy. Ampicillin alone or with clavulanic acid may be used, but many strains of *E. coli* and other organisms

implicated in traveler's diarrhea are resistant. Azithromycin or an oral third-generation cephalosporin may be the best option.

Because of the major problems encountered when infants are given local foods and beverages, women are strongly encouraged to breast-feed when traveling with a neonate. A nursing mother with traveler's diarrhea should not stop breast-feeding but should increase her fluid intake.

Air Travel and High-Altitude Destinations Commercial air travel is not a risk to the healthy pregnant woman or to the fetus. Fetal oxygenation is not adversely affected by the decreased cabin pressures because of the fetal hemoglobin dissociation curve; the higher radiation levels reported at altitudes>10,500 m (35,000 ft) should pose no problem to the healthy pregnant traveler. Since each airline has a policy regarding pregnancy and flying, it is best to check with the specific carrier when booking reservations. Domestic air travel is usually permitted until the 36th week, whereas international air travel is generally curtailed after the 32nd week.

There are no known risks for pregnant women who travel to high-altitude destinations and stay for short periods. However, there are likewise no data on the safety of pregnant women at altitudes>4500 m (15,000 ft). Because of the harsh conditions usually associated with such trips, they are generally contraindicated for other reasons.

THE HIV-INFECTED TRAVELER

The traveler infected with HIV is at special risk of serious infections due to a number of pathogens that may be more prevalent at travel destinations than at home. However, the degree of risk depends primarily on the state of the immune system at the time of travel. For persons whose CD4+ cell counts are normal or>500/uL, no data suggest a greater risk during travel than for persons without HIV infection. Individuals with AIDS (CD4+ counts <200/uL) and others who are symptomatic need special counseling and should visit a travel medicine practitioner before departure, especially when traveling to the developing world.

Several countries now routinely deny entry to HIV-positive individuals, even though no data show that these restrictions decrease rates of transmission of the virus. In general, HIV testing is required of those individuals who wish to stay abroad longer than 3 months or who intend to work or study abroad. Some countries will accept an HIV serologic test done within 6 months of departure, whereas others will not accept a blood test done at any time in the traveler's home country. In addition, border officials often have the authority to make inquiries of individuals entering a country and to check the medications they are carrying. If a drug such as zidovudine (AZT) is identified, the person may be barred from entering the country. Information on testing requirements for specific countries is available from consular offices but is subject to frequent change.

Health insurance policies should be checked to make sure they are valid for care in other countries. The HIV-positive traveler should strongly consider obtaining trip cancellation insurance and evacuation insurance in case of illness. It is ideal to have the name of a physician at the travel destination who is familiar with the treatment of patients with AIDS, as the clinical findings associated with infection may be atypical in a

patient with AIDS, and several infections may exist simultaneously. The traveler should be encouraged to visit the physician promptly if problems arise.

Immunizations All of the HIV-infected traveler's routine immunizations should be up to date (Chap. 122). The response to immunization may be impaired at CD4+ cell counts of <200/uL (and in some cases at even higher counts). However, when the risk of illness is high or the sequelae of illness are serious, immunization is recommended. In certain circumstances, it may be prudent to check the adequacy of the serum antibody response before departure (e.g., yellow fever neutralization inhibition if exposure is unavoidable).

Because of the increased risk of infections due to *Streptococcus pneumoniae* and other bacterial pathogens that cause pneumonia following influenza, pneumococcal polysaccharide and influenza vaccines should be administered. The estimated rates of response to influenza vaccine are >80% among persons with asymptomatic HIV infection and<50% among those with AIDS.

In general, live attenuated vaccines are contraindicated for persons with immune dysfunction. Live oral polio vaccine should not be given to HIV-infected patients or to members of their households. Instead, inactivated polio vaccine (eIPV) should be used; most HIV-infected individuals without AIDS will develop protective antibody levels in response to this vaccine.

Because measles (rubeola) can be a severe and lethal infection in HIV-positive patients, the measles vaccine (or the combination measles-mumps-rubella vaccine) should be given to these individuals. Although this is a live vaccine, there have been no reports of serious complications in this population. Between 18 and 58% of symptomatic HIV-infected vaccinees develop adequate antibody titers, and between 50 and 100% of those who are infected but asymptomatic seroconvert.

The decision of whether or not to administer any of the special vaccines to an HIV-infected traveler should be based on the individual's risk. Inactivated vaccines can be administered without concern for safety but with concern about adequate protection. For example, data suggest that HIV-infected persons do not have as strong an antibody response to the meningococcal meningitis vaccine as do uninfected persons. Moreover, few data are available on the efficacy of many of the other vaccines (e.g., those for hepatitis A, typhoid, and cholera).

It is recommended that the live yellow fever vaccine not be given to HIV-infected travelers. Nevertheless, when inadvertently administered to HIV-positive military personnel, this vaccine elicited no adverse reactions. Therefore, if the traveler's CD4+ count is >200/uL and travel in an endemic area is absolutely necessary, the vaccine can probably be administered safely. HIV-infected persons whose CD4+ count is<200/uL should be discouraged from traveling to endemic regions. If the traveler is passing through or traveling to an area where the vaccine is required but the disease risk is low, a physician's waiver should be issued. Bacille Calmette-Guerin vaccine should not be given because of reports of disseminated infection in HIV-infected persons.

A transient (days to weeks) burst of viremia has been demonstrated in HIV-infected

individuals following immunization with vaccines for such diseases as influenza, pneumococcal infection, and tetanus (Chap. 309). However, at this point, there is no evidence that this transient increase in viremia is detrimental over time. Furthermore, it is likely that immune activation associated with infection with the live organisms in question would result in increases in viremia of greater magnitude and duration than those associated with vaccination. Therefore, the vaccination recommendations discussed above need not be modified at this time.

Gastrointestinal Illness Decreased levels of gastric acid, abnormal gastrointestinal mucosal immunity, other complications of HIV infection, and medications taken by HIV-infected patients make traveler's diarrhea especially problematic in these individuals. Traveler's diarrhea is likely to occur more frequently, be more severe, and be more difficult to treat in association with HIV infection. *Salmonella*, *Shigella*, and *Campylobacter* infections are also more protracted and more often accompanied by bacteremia in HIV-infected persons.

Cryptosporidium (Chap. 218), a common cause of diarrhea in tropical countries, produces severe chronic diarrhea and cholecystitis with increased mortality among patients with AIDS. *Isospora belli* causes infections at high rates among AIDS patients in the developing world; this infection is associated with malabsorption, weight loss, and relapses after treatment. Persistent diarrhea due to microsporidiosis has been reported.

Because of these potential problems, the HIV-infected traveler must be careful to consume only appropriately prepared foods and beverages. In addition, this group of individuals may benefit from prophylaxis for traveler's diarrhea, using bismuth subsalicylate or a daily antibiotic (ideally a quinolone derivative) for short-term travel to the developing world. If the traveler is already taking a sulfonamide preparation for prophylaxis of *Pneumocystis* pneumonia, a regimen of self-treatment with a quinolone would be appropriate.

Other Travel-Related Infections Data are lacking on the severity of vector-borne diseases in HIV-infected individuals. Malaria is especially severe in asplenic and certain immunocompromised hosts, although increased severity has not been demonstrated in AIDS. *Babesia* infection is known to cause serious illness and to recur in HIV-infected patients; this tick-transmitted illness occurs in parts of the United States but is not known to be a widespread problem.

Visceral leishmaniasis (Chap. 215) has been reported in numerous HIV-infected travelers. Because the usual signs -- splenomegaly and hyperglobulinemia -- are nonspecific and may even be lacking, the diagnosis is difficult to make. In addition, serologic results are often negative. This infection is difficult to treat, and its associated mortality is high. Even short-term travelers to southern Europe have developed the illness; thus, the avoidance of sandfly bites is critical.

Certain respiratory illnesses, such as histoplasmosis and coccidioidomycosis, cause greater morbidity and mortality among patients with AIDS than in the general population. Though tuberculosis is common among HIV-infected persons (especially in developing countries), the acquisition of this infection by the short-term traveler is not a major concern. The possibility of acquiring *Legionella* infections from spas should be

considered, although no data confirm an increase in the severity of such infections in AIDS.

Finally, the HIV-infected traveler should always be cautioned about safe sexual practices, which may help prevent both the transmission of HIV to others and the acquisition by the traveler of other sexually transmitted diseases that may be drug resistant or may result in serious sequelae (e.g., syphilis).

Medications Adverse events due to medications and drug interactions are common and raise complex issues for HIV-infected persons. In addition, rates of cutaneous reaction are unusually high among patients with AIDS. Physicians advising these travelers need to consider the problems that may arise from the use of agents such as antimalarial drugs, medications for altitude acclimatization, or antidiarrheal compounds; one example is increased cutaneous sensitivity to sulfonamides. Since zidovudine is metabolized by hepatic glucuronidation, inhibitors of this process may elevate serum levels of the drug. Though quinine does not affect levels of zidovudine, there are no relevant data on chloroquine, primaquine, or mefloquine. Furthermore, it is not known whether the antagonistic effect of zidovudine on pyrimethamine has clinical relevance in the treatment or prevention of plasmodial infections.

CHRONIC ILLNESS, DISABILITY, AND TRAVEL

Evaluating fitness for travel is a growing issue in view of the increased number of elderly and chronically ill individuals journeying to exotic destinations. Conditions encountered during flight are of particular concern in these cases. Since most commercial aircraft are pressurized to 2500 m (8000 ft) above sea level, corresponding to a Pa_{O_2} of about 55 mmHg, individuals with serious cardiopulmonary problems should be evaluated before travel. In addition, those who have recently had surgery, a myocardial infarction, a cerebrovascular accident, or another medical crisis may be at high risk for adverse events in flight. A summary of current recommendations regarding fitness to fly has been published by the Aerospace Medical Association Air Transport Medical Committee. Chronic health problems should not prevent travel, but special measures can make the journey safer and more comfortable.

Heart Disease Cardiovascular events are the main cause of deaths among travelers and of in-flight emergencies on commercial aircraft. Persons with underlying heart disease should review their itineraries with a physician prior to departure; travel in harsh environments or to remote destinations is not wise. Extra supplies of all medications should be kept in carry-on luggage, along with a recent copy of an electrocardiogram and the name and telephone number of the traveler's physician at home. Pacemakers are not affected by airport security devices, but electronic telephone checks of pacemaker function cannot be transmitted by international satellites. The traveler may benefit from supplemental oxygen, which should be ordered by a physician (since oxygen delivery systems are not standard) 48 to 72 h before flight time. Personal oxygen tanks are not permitted on aircraft. Travelers should request aisle seating and should walk, perform stretching and flexing exercises, and remain hydrated during the flight to prevent venous thrombosis and pulmonary embolism.

Chronic Lung Disease Chronic obstructive pulmonary disease (COPD) is one of the

most common diagnoses in patients who require emergency-room evaluation for symptoms occurring during airline flights. Patients with COPD experience dyspnea, edema, wheezing, cyanosis, and chest pain. The best predictor of the development of these symptoms is the sea level Pa_{O2}. A Pa_{O2} of at least 72 mmHg corresponds to an in-flight Pa_{O2} of 55 mmHg when the cabin is pressurized to 2500 m (8000 ft). Therefore, if the traveler's baseline Pa_{O2} is<72 mmHg, the provision of supplemental oxygen during the flight should be considered. Pulmonary function is also maximized by continuing bronchodilator treatment and the use of glucocorticoids as prescribed. Contraindications to flight include active bronchospasm, lower respiratory infection, phlebitis, pulmonary hypertension, and recent thoracic surgery (within the preceding 3 weeks) or pneumothorax. Consideration should be given to decreasing the amount of outdoor activity at the destination if there is excessive air pollution.

Diabetes Mellitus Alterations in glucose control and changes in insulin requirements are common problems when diabetic patients travel. Changes in time zone, in the amount and timing of food intake, and in physical activity demand more vigilant assessment of metabolic control. The diabetic traveler should pack medication (including a bottle of regular insulin for emergencies), insulin syringes and needles, equipment and supplies for glucose monitoring, and snacks in carry-on luggage. Insulin is stable for about 3 months at room temperature but should be kept as cool as possible. The name and telephone number of the home physician and a card and necklace listing the medical problems and the type and dose of insulin used should accompany the traveler. When six or more time zones are crossed, insulin requirements may be temporarily altered, depending on food intake and physical activity. In traveling eastward (e.g., from the United States to Europe), the morning insulin dose on arrival may need to be decreased. The blood glucose can then be checked during the day to determine whether additional insulin is required. For flights westward, with lengthening of the day, an additional dose of regular insulin may be required. Comfortable footwear is essential for the diabetic traveler.

Other Special Groups Other groups for whom special travel measures are now being encouraged include patients undergoing dialysis, those with transplants, and those with other disabilities. Up to 13% of travelers have some disability, but few advocacy groups and tour companies dedicate themselves to this growing population. The key to safe travel in each case is adequate research ahead of time. Patients undergoing chronic ambulatory peritoneal dialysis may ship their dialysis solutions to their destinations before traveling. They should carry essential medical records as well as antibiotics for self-treatment of presumed peritonitis. Hemodialysis patients need to reserve appointments at dialysis centers prior to their departure from home. Travel by transplant recipients to distant destinations should ideally be scheduled at least 1 year after surgery, as most rejection episodes occur early. Medication interactions are a source of serious concern for these travelers, and appropriate medical information should be carried, along with the home physician's name and telephone number. Some travelers taking glucocorticoids carry stress doses in case they become ill. Immunization of these immunocompromised travelers may result in less than adequate protection against certain diseases. Thus, the traveler and physician must carefully consider which destinations are appropriate.

PROBLEMS AFTER RETURN

The most frequent medical problems encountered by travelers after their return home are diarrhea, fever, respiratory illnesses, and skin diseases. Frequently ignored problems are fatigue and emotional stress, especially in long-stay travelers. The approach to diagnosis requires some knowledge of geographic medicine, in particular the epidemiology and clinical presentation of infectious disorders. A geographic history should focus on the traveler's exact itinerary, including dates of arrival and departure; exposure history (food indiscretions, drinking-water sources, freshwater contact, sexual activity, animal contact, insect bites); location and style of travel (urban vs. rural, first-class hotel accommodation vs. camping); immunization history; and use of antimalarial chemosuppression.

DIARRHEA

Although extremely common, acute traveler's diarrhea is usually self-limited or amenable to antibiotic therapy. Bowel symptoms that persist after the traveler's return home have a less well-defined etiology and may require medical attention from a specialist. Infectious agents appear to be responsible for only a small proportion of cases with persistent bowel symptoms. Of the pathogens detected in these instances, *Giardia lamblia* (Chap. 218) is by far the most common; *Cyclospora cayetanensis*, *Cryptosporidium* spp., and *Entamoeba histolytica* are rare isolates. The most frequent causes of persistent diarrhea after travel are postinfectious sequelae, such as lactose intolerance or an irritable bowel syndrome. When no infectious etiology can be identified, a trial of metronidazole therapy for presumed giardiasis, a strict lactose-free diet for 1 week, or a several-week trial of high-dose hydrophilic mucilloid relieves the symptoms of many patients.

FEVER

Fever in a traveler who has returned from a malarious area should be considered a medical emergency because death from *P. falciparum* malaria can follow an illness of only several days' duration. Although "fever from the tropics" does not always have a tropical cause, malaria should be the first diagnosis considered. The risk of *P. falciparum* malaria is highest among travelers returning from Africa or Oceania and among those who become symptomatic within the first 2 months after return. Other important causes of fever after travel include viral hepatitis (hepatitis A and E), typhoid fever, bacterial enteritis, arbovirus infections (e.g., dengue fever), rickettsial infections (including tick and scrub typhus or Q fever) and -- in rare instances -- leptospirosis, acute HIV infection, and amebic liver abscess. In at least 25% of cases, no etiology can be found, and the illness resolves spontaneously. Clinicians should keep in mind that no present-day antimalarial agent guarantees protection from malaria and that some immunizations -- notably, those against typhoid and cholera -- are only partially protective.

As noted above, the approach to the febrile returned traveler begins with a detailed medical and geographic history. Knowing exact dates of arrival and departure from tropical areas enables the physician to ascertain the shortest and longest possible incubation periods for illnesses in the differential diagnosis. For example, a traveler who develops fever <1 week after arrival in a malarious area cannot have malaria because

the incubation period is too short, whereas a fever whose onset comes >2 weeks after departure from an endemic area cannot be dengue fever because the incubation period is too long. In the physical examination, particular attention should be given to the skin so as not to miss a subtle rash or eschar.

When no specific diagnosis is forthcoming, the following investigations may be helpful: complete blood count, liver function tests, thick/thin blood films for malaria (repeated twice if necessary), urinalysis, blood cultures (repeated once if necessary), and collection of an acute-phase serum sample to be held for later examination along with a paired convalescent-phase serum sample.

SKIN DISEASES

Pyodermas, sunburn, insect bites, skin ulcers, and cutaneous larva migrans are the most common skin conditions encountered in travelers after their return home. In those with persistent skin ulcers, the diagnoses of cutaneous leishmaniasis, mycobacterial infection, or fungal infection should be considered. Careful, complete inspection of the skin is important in detecting the rickettsial eschar in a febrile patient or the central breathing hole in a "boil" due to myiasis.

EMERGING INFECTIOUS DISEASES

In recent years, travel and commerce have fostered the worldwide spread of HIV infection, led to the reemergence of cholera as a global health threat, and created considerable fear about the possible spread of Ebola virus infection and plague. For travelers, there are more realistic concerns. One of the largest outbreaks of dengue fever ever documented is now raging in Latin America; schistosomiasis is being described in previously unaffected lakes in Africa; and antibiotic-resistant strains of sexually transmitted and enteric pathogens are emerging at an alarming rate in the developing world. As Nobel Laureate Dr. Joshua Lederberg pointed out, "The microbe that felled one child in a distant continent yesterday can reach yours today and seed a global pandemic tomorrow." The vigilant clinician understands that the importance of a thorough travel history cannot be overemphasized.

SOURCES OF INFORMATION ON TRAVEL MEDICINE

· CDC publication *Health Information for International Travel*

· CDC home page:http://www.cdc.gov

·CDC travel information:http://www.cdc.gov/travel/travel.html

·Health Canada:http://www.hwc.ca/hpb/lcdc

·International Society of Travel Medicine:http://www.istm.org

(Bibliography omitted in Palm version)

Back to Table of Contents

SECTION 2 -CLINICAL SYNDROMES: COMMUNITY-ACQUIRED INFECTIONS

124. SEPSIS AND SEPTIC SHOCK - *Robert S. Munford*

DEFINITIONS (See Table 124-1)

The host's reaction to invading microbes involves a rapidly amplifying polyphony of signals and responses that may spread beyond the invaded tissue. Fever or hypothermia, tachypnea, and tachycardia often herald the onset of *sepsis*, the systemic response to microbial invasion. When counterregulatory control mechanisms are overwhelmed, homeostasis may fail, and dysfunction of major organs may supervene (*severe sepsis*). Further regulatory imbalance leads to *septic shock*, which is characterized by hypotension as well as organ dysfunction. As sepsis progresses to septic shock, the risk of dying increases substantially. Sepsis is usually reversible, whereas patients with septic shock often succumb despite aggressive therapy.

The *systemic inflammatory response syndrome* (SIRS), as defined in the early 1990s by critical care specialists, may have an infectious or a noninfectious etiology. If infection is suspected or proven, a patient with SIRS is said to have sepsis.

ETIOLOGY

Sepsis can be a response to any class of microorganism. Microbial invasion of the bloodstream is not essential for the development of sepsis, since local or systemic spread of microbial signal molecules or toxins can also elicit the response. Blood cultures yield bacteria or fungi in ~20 to 40% of cases of severe sepsis and 40 to 70% of cases of septic shock. Individual gram-negative or gram-positive bacteria account for ~70% of these isolates; the remainder are fungi or a mixture of microorganisms (Table 124-2). In patients whose blood cultures are negative, the etiologic agent is often established by culture or microscopic examination of infected material from a local site. In some case series, a majority of patients with a clinical picture of severe sepsis or septic shock have had negative microbiologic data.

Factors that predispose to gram-negative bacillary bacteremia include diabetes mellitus, lymphoproliferative diseases, cirrhosis of the liver, burns, invasive procedures or devices, and treatment with drugs that cause neutropenia. Major risk factors for gram-positive bacteremia include vascular catheterization, the presence of indwelling mechanical devices, burns, and intravenous drug use. Fungemia occurs most often in immunosuppressed patients with neutropenia, often after broad-spectrum antimicrobial therapy. In patients who experience bacteremia, factors that increase the risk of developing severe sepsis include age (>50 years) and a primary pulmonary, abdominal, or neuromeningeal site of infection. Bacteremia that arises from an intravascular catheter or the urinary tract is less likely to induce severe sepsis.

EPIDEMIOLOGY

The septic response is now a contributing factor in>100,000 deaths per year in the United States. The incidence of severe sepsis and septic shock has increased over the last 15 years and is now probably between 300,000 and 500,000 cases per year.

Approximately two-thirds of cases occur in patients hospitalized for other illnesses. The increasing incidence of severe sepsis in the United States is attributable to the aging of the population, the increasing longevity of patients with chronic diseases, and the relatively high frequency with which sepsis develops in patients with AIDS. The widespread use of antimicrobial agents, glucocorticoids, indwelling catheters and mechanical devices, and mechanical ventilation also plays a role.

PATHOPHYSIOLOGY

The septic response is often triggered when microorganisms spread from the gastrointestinal tract or skin into contiguous tissues. Localized tissue infection may then lead to bacteremia or fungemia. Alternatively, microorganisms may be introduced directly into the bloodstream (for example, via intravenous catheters). In general, the septic response occurs when immune defenses fail to contain an invading microbe. Since most cases are triggered by microbes that do not ordinarily cause systemic disease in normal hosts (Table 124-2), deficiencies in nonadaptive host factors may be most important. The septic response may also be induced by microbial exotoxins that act as superantigens (e.g., toxic shock syndrome toxin 1;Chap. 139).

Microbial Signals Animals recognize certain microbial molecules as signals that microorganisms have invaded. Lipopolysaccharide (LPS, also called *endotoxin*) is the most potent and best-studied gram-negative bacterial signal molecule. A plasma protein (LPS-binding protein, or LBP) transfers LPS to CD14 on the surfaces of monocytes, macrophages, and neutrophils. This interaction rapidly triggers the production and release of mediators, such as tumor necrosis factor (TNF) a (see below), that amplify the LPS signal and transmit it to other cells and tissues. Soluble CD14 may also bind LPS in plasma and transfer it to cells that lack cell-surface CD14. The peptidoglycan and lipoteichoic acids of gram-positive bacteria, certain polysaccharides, extracellular enzymes, and toxins elicit responses in animals that are similar to those induced by LPS; some of these molecules may also bind CD14. CD14 thus attracts numerous non-self molecules to the surfaces of myeloid cells, greatly increasing the sensitivity with which these molecules can be recognized by the host. Evidence suggests that signal specificity occurs in or on the plasma membrane and is conferred, at least in part, by members of the toll-like receptor family of transmembrane proteins. Other innate immune mechanisms for recognizing microbial molecules include complement (principally the alternative pathway), mannose-binding protein, and C-reactive protein.

Host Responses The septic response involves complex interactions among microbial signal molecules, leukocytes, humoral mediators, and the vascular endothelium.

Cytokines Inflammatory cytokines amplify and diversify the response. These proteins can exert endocrine, paracrine, and autocrine effects (Chap. 305).TNF-astimulates leukocytes and vascular endothelial cells to release other cytokines (as well as additional TNF-a), to express cell-surface adhesion molecules, and to increase arachidonic acid turnover. Blood levels of TNF-a are high in most patients with severe sepsis or septic shock. Moreover, intravenous infusion of TNF-acan elicit many of the characteristic abnormalities of sepsis, including fever, tachycardia, tachypnea, leukocytosis, myalgias, and somnolence. In animals, larger doses of TNF-a induce shock, disseminated intravascular coagulation (DIC), and death. Specific TNF-a

antagonists can abrogate the septic response and prevent the deaths of experimental animals challenged with endotoxin.

AlthoughTNF-ais a central mediator, it is only one of many cytokines that contribute to the septic process. Interleukin (IL) 1b, for example, which exhibits many of the same activities as TNF-a, seems to play an increasingly significant role as sepsis intensifies. TNF-a, IL-1b, interferon-g, IL-8, and other cytokines probably interact synergistically with each other and with additional mediators. Moreover, some mediators (such as IL-1b and TNF-a) may enhance their own rates of synthesis by positive feedback. As sepsis progresses, the mixture of cytokines and other molecules becomes very complex: elevated blood levels of>50 molecules have been found in patients with septic shock. In animal models, the septic response can be interrupted by early interventions that neutralize one or another of its many components; this observation testifies to the importance of mediator interactions for the overall outcome. It has been much more difficult to rescue animals from severe sepsis and septic shock.

Phospholipid-Derived Mediators Arachidonic acid, released from membrane phospholipids by phospholipase A_2, is converted by the cyclooxygenase pathway into prostaglandins and thromboxanes. Prostaglandin E_2 and prostacyclin cause peripheral vasodilatation, whereas thromboxane is a vasoconstrictor and promotes platelet aggregation. Administration of the cyclooxygenase inhibitor ibuprofen for 48 h to patients with severe sepsis suppressed production of these metabolites and decreased body temperature, heart rate, and metabolic acidosis without reducing mortality. Leukotrienes are also potent mediators of ischemia and shock; the fact that the reaction to endotoxin challenge is normal in mice that lack the 5-lipoxygenase gene, however, casts doubt on the role of leukotrienes in the septic response.

Another important phospholipid-derived mediator is platelet-activating factor (PAF; 1-*O*-alkyl-2-acetyl-*sn*-glycero-3-phosphorylcholine). PAF potently stimulates neutrophil aggregation and degranulation, promotes platelet aggregation, and may contribute to tissue injury.

Coagulation Factors Intravascular fibrin deposition, thrombosis, andDIC are important features of the septic response.IL-6 and other mediators promote intravascular coagulation initially by inducing blood monocytes to express tissue factor (Chap. 62). When tissue factor is expressed on monocytes, it binds to factor VIIa to form an active complex that can convert factors X and IX to enzymatically active forms. The result is activation of both extrinsic and intrinsic clotting pathways, culminating in the generation of fibrin. Clotting is also favored by impaired function of the protein C-protein S inhibitory pathway and depletion of antithrombin, while fibrinolysis is prevented by increased plasma levels of plasminogen activator inhibitor 1. Thus, there may be a striking propensity to intravascular fibrin deposition, thrombosis, and bleeding (Chap. 146). Contact-system activation occurs during sepsis but contributes more to the development of hypotension than to DIC.

Complement C5a and other products of complement activation may promote neutrophil reactions such as chemotaxis, aggregation, degranulation, and oxygen-radical production. When administered to animals, C5a induces hypotension, pulmonary vasoconstriction, neutropenia, and vascular leakiness due in part to endothelial

damage.

Activation of the Vascular Endothelium Many tissues may be damaged by the septic response. The probable underlying mechanism is widespread vascular endothelial injury, with fluid extravasation and microthrombosis that decrease oxygen and substrate utilization by the affected tissues. Leukocyte-derived mediators and platelet-leukocyte-fibrin thrombi contribute to this injury, but the vascular endothelium itself seems to play an active role. Stimuli such asTNF-ainduce vascular endothelial cells to produce and release cytokines, procoagulant molecules,PAF, endothelium-derived relaxing factor (nitric oxide), and other mediators. In addition, regulated cell-adhesion molecules promote the adherence of neutrophils to endothelial cells. While these responses may attract phagocytes to infected sites and activate their antimicrobial arsenals, endothelial cell activation can also promote increased vascular permeability, microvascular thrombosis,DIC, and hypotension. Moreover, vascular integrity may be damaged by neutrophil enzymes (such as elastase) and toxic oxygen metabolites so that local hemorrhage ensues. Blocking the adhesion of leukocytes to endothelial cell surfaces, as with monoclonal antibodies to intercellular adhesion molecule 1, can prevent tissue necrosis in response to endotoxin administration in animals.

Septic Shock Much evidence now implicates nitric oxide, produced by inducible nitric oxide synthase (iNOS), as a mediator of septic shock in experimental animals and probably in humans. Mice that lack the *iNOS* gene may not be resistant to endotoxic shock, however. Other prominent hypotensive molecules are b-endorphin, bradykinin,PAF, and prostacyclin. Agents that inhibit the synthesis or action of each of these mediators can prevent or reverse endotoxic shock in animals. However, in clinical trials, neither a PAF receptor antagonist nor a bradykinin antagonist improved the survival rate of patients with septic shock, and a NOS inhibitor, L-N_G-methylarginine HCl, actually increased the mortality rate.

Control Mechanisms Elaborate host mechanisms regulate both microbial signals and the inflammatory response. While plasma LBP promotes the inflammatory response by facilitating the interaction ofLPS with monocyte cell-surface CD14, LBP and other plasma proteins (such as phospholipid transfer protein) also can prevent LPS signaling by transferring LPS molecules into plasma lipoprotein particles. The relative plasma concentrations of LPS, LBP, CD14, and lipoproteins may therefore govern the intensity with which LPS -- and probably other microbial molecules -- can trigger host responses. The mechanisms that control the inflammatory response are also complex, overlapping, and poorly understood. Glucocorticoids inhibit cytokine synthesis by monocytes in vitro and, when administered with or shortly after an inflammatory stimulus, may protect animals from septic shock. The increase in blood cortisol levels early in the septic response presumably plays a similar inhibitory role. In addition, certain cytokine antagonists may contribute. Blood levels ofIL-1 receptor antagonist (IL-1Ra) often greatly exceed those of circulating IL-1b, and this excess may result in inhibition of the binding of IL-1b to its receptors. Transforming growth factor b (TGF-b) and IL-10 can also inhibit LPS-induced responses by human monocytes in vitro and prevent endotoxic death in animals. Blood and tissue levels of prostaglandin E_2, TGF-b,a-melanocyte-stimulating hormone, cortisol, IL-1Ra, soluble TNF receptors, and IL-10 increase during the septic response, and these molecules probably act in concert

to diminish its intensity. In fact, very high concentrations of many anti-inflammatory molecules are found in the blood of patients with severe sepsis or septic shock, so that the net mediator balance in the blood of these extremely sick patients may actually be anti-inflammatory. In addition, blood leukocytes from patients with severe sepsis are often hyporesponsive to agonists such as LPS. In patients with severe sepsis, persistence of leukocyte hyporesponsiveness has been associated with an increased risk of dying. Research is needed to clarify the role of the anti-inflammatory response in the septic process.

CLINICAL MANIFESTATIONS

The manifestations of the septic response are usually superimposed on the symptoms and signs of the patient's underlying illness and primary infection. The systemic response to infection often intensifies over time from mild (sepsis) to extremely severe (septic shock). The rate at which the response increases may differ from patient to patient, and there are striking individual variations in its manifestations. For example, some patients with sepsis are normo- or hypothermic; the absence of fever is most common in neonates, in elderly patients, and in persons with uremia or alcoholism.

Hyperventilation is often an early sign. Disorientation, confusion, and other manifestations of encephalopathy may also develop early in the septic response, particularly in the elderly and in individuals with preexisting neurologic impairment. Focal neurologic signs are uncommon, although preexisting focal deficits may become more prominent.

Hypotension andDICpredispose to acrocyanosis and ischemic necrosis of peripheral tissues, most commonly the digits (Fig. 124-CD1). Cellulitis, pustules, bullae, or hemorrhagic lesions may develop when hematogenous bacteria or fungi seed the skin or underlying soft tissue (Fig. 124-CD2). Bacterial toxins may also be distributed hematogenously to elicit diffuse cutaneous reactions. On occasion, skin lesions may suggest specific pathogens. When sepsis is accompanied by cutaneous petechiae or purpura, infection with *Neisseria meningitidis* (or, less commonly, *Haemophilus influenzae*) should be suspected (Plate IID-44); in a patient who has been bitten by a tick while in an endemic area, petechial lesions also suggest Rocky Mountain spotted fever (Plate IID-45). A cutaneous lesion seen almost exclusively in neutropenic patients is ecthyma gangrenosum (Fig. 19-CD1), usually caused by *Pseudomonas aeruginosa*. It is a bullous lesion, surrounded by edema, that undergoes central hemorrhage and necrosis (Plate IID-57C). Histopathologic examination shows bacteria in and around the wall of a small vessel, with little or no neutrophilic response. Hemorrhagic or bullous lesions in a septic patient who has recently eaten raw oysters suggest *Vibrio vulnificus* bacteremia, while such lesions in a patient who has recently suffered a dog bite may indicate bloodstream infection due to *Capnocytophaga canimorsus* or *C. cynodegmi*. Generalized erythroderma in a septic patient suggests the toxic shock syndrome due to *Staphylococcus aureus* or *Streptococcus pyogenes*.

Gastrointestinal manifestations such as nausea, vomiting, diarrhea, and ileus may suggest acute gastroenteritis. Stress ulceration can lead to upper gastrointestinal bleeding. Cholestatic jaundice, with elevated levels of serum bilirubin (mostly conjugated) and alkaline phosphatase, may precede other signs of sepsis.

Hepatocellular or canalicular dysfunction appears to underlie most cases, and the results of hepatic function tests return to normal with resolution of the infection. Prolonged or severe hypotension may induce acute hepatic injury or ischemic bowel necrosis.

Many tissues may be unable to extract oxygen normally from the blood, so that anaerobic metabolism occurs despite near-normal mixed venous oxygen saturation. Blood lactate levels rise early, in part because of increased glycolysis with impaired clearance of the resulting lactate and pyruvate by the liver and kidneys. As hypoperfusion develops, tissue hypoxia generates more lactic acid, contributing to metabolic acidosis. The blood glucose concentration often increases, particularly in patients with diabetes, although impaired gluconeogenesis and excessive insulin release on occasion produce hypoglycemia. The cytokine-driven acute-phase response inhibits the synthesis of albumin and transthyretin while enhancing the production of C-reactive protein,LBP, fibrinogen, and complement components. Protein catabolism is often markedly accelerated.

MAJOR COMPLICATIONS

Cardiopulmonary Complications Ventilation-perfusion mismatching produces a fall in arterial P_{O_2}early in the course. Increasing alveolar capillary permeability results in an increased pulmonary water content, which decreases pulmonary compliance and interferes with oxygen exchange. Progressive diffuse pulmonary infiltrates and arterial hypoxemia (Pa_{O_2}/FI_{O_2},<200 mmHg) indicate the development of the acute respiratory distress syndrome (ARDS). ARDS develops in ~50% of patients with severe sepsis or septic shock. The failure of the respiratory muscles can exacerbate hypoxemia and hypercapnia. An elevated pulmonary capillary wedge pressure (>18 mmHg) suggests fluid volume overload or cardiac failure rather than ARDS. Pneumonia caused by viruses or by *Pneumocystis carinii* may be clinically indistinguishable from ARDS.

Sepsis-induced hypotension usually results from a generalized maldistribution of blood flow and blood volume and from hypovolemia that is due, at least in part, to diffuse capillary leakage of intravascular fluid. Other factors that may decrease effective intravascular volume include dehydration from antecedent disease or insensible fluid losses, vomiting or diarrhea, and polyuria. During early septic shock, systemic vascular resistance is usually elevated and cardiac output may be low. After fluid repletion, in contrast, cardiac output typically increases and systemic vascular resistance falls. Indeed, normal or increased cardiac output and decreased systemic vascular resistance distinguish septic shock from cardiogenic, extracardiac obstructive, and hypovolemic shock; other processes that can produce this combination include anaphylaxis, beriberi, cirrhosis, and overdoses of nitroprusside or narcotics (Chap. 38).

Depression of myocardial function, manifested as increased end-diastolic and systolic ventricular volumes with a decreased ejection fraction, develops within 24 h in most patients with severe sepsis. Cardiac output is maintained despite the low ejection fraction because ventricular dilatation permits a normal stroke volume. In survivors, myocardial function returns to normal over several days. Although myocardial dysfunction may contribute to hypotension, refractory hypotension is usually due to a low systemic vascular resistance, and death results from refractory shock or the failure

of multiple organs rather than from cardiac dysfunction per se.

Renal Complications Oliguria, azotemia, proteinuria, and nonspecific urinary casts are frequently found. Many patients are inappropriately polyuric; hyperglycemia may exacerbate this tendency. Most renal failure is due to acute tubular necrosis induced by hypotension or capillary injury, although some patients also have glomerulonephritis, renal cortical necrosis, or interstitial nephritis. Drug-induced renal damage may complicate therapy, particularly when hypotensive patients are given aminoglycoside antibiotics.

Coagulation Although thrombocytopenia occurs in 10 to 30% of patients, the underlying mechanism(s) are not understood. Platelet counts are usually very low (<50,000/uL) in patients withDIC; these low counts typically reflect diffuse endothelial injury or microvascular thrombosis.

Neurologic Complications When the septic illness lasts for weeks to months, "critical-illness" polyneuropathy may prevent weaning from ventilatory support and produce distal motor weakness. Electrophysiologic studies are diagnostic. Guillain-Barre syndrome, metabolic disturbances, and toxin activity must be ruled out.

LABORATORY FINDINGS

Abnormalities that occur early in the septic response may include leukocytosis with a left shift, thrombocytopenia, hyperbilirubinemia, and proteinuria. Leukopenia may develop. The neutrophils may contain toxic granulations (Fig. 124-CD3), Dohle bodies, or cytoplasmic vacuoles. As the septic response becomes more severe, thrombocytopenia worsens (often with prolongation of the thrombin time, decreased fibrinogen, and the presence of D-dimers, suggestingDIC), azotemia and hyperbilirubinemia become more prominent, and levels of aminotransferases rise. Active hemolysis suggests clostridial bacteremia, malaria, a drug reaction, or DIC; in the case of DIC, microangiopathic changes may be seen on a blood smear.

During early sepsis, hyperventilation induces respiratory alkalosis. With respiratory muscle fatigue and the accumulation of lactate, metabolic acidosis (with increased anion gap) typically supervenes. Evaluation of arterial blood gases reveals hypoxemia, which is initially correctable with supplemental oxygen but whose later refractoriness to 100% oxygen inhalation indicates right-to-left shunting. The chest radiograph may be normal or may show evidence of underlying pneumonia, volume overload, or the diffuse infiltrates ofARDS. The electrocardiogram may show only sinus tachycardia or nonspecific ST-T wave abnormalities.

Most diabetic patients with sepsis develop hyperglycemia. Severe infection may precipitate diabetic ketoacidosis, which may exacerbate hypotension (Chap. 333). Hypoglycemia occurs rarely. The serum albumin level, initially within the normal range, declines as sepsis continues. Serum lipid concentrations are often elevated. Hypocalcemia is rare.

DIAGNOSIS

There is no specific test. Diagnostically sensitive findings in a patient with suspected or proven infection include fever or hypothermia, tachypnea, tachycardia, and leukocytosis or leukopenia (Table 124-1); acutely altered mental status, thrombocytopenia, or hypotension also suggests the diagnosis. The septic response can be quite variable, however. In one study, 36% of patients with severe sepsis had a normal temperature, 40% had a normal respiratory rate, 10% had a normal pulse rate, and 33% had normal white blood cell counts. Moreover, the systemic responses of uninfected patients with other conditions may be similar to those characteristic of sepsis. Noninfectious etiologies ofSIRS(Table 124-1) include pancreatitis, burns, trauma, adrenal insufficiency, pulmonary embolism, dissecting or ruptured aortic aneurysm, myocardial infarction, occult hemorrhage, cardiac tamponade, post-cardiopulmonary bypass syndrome, anaphylaxis, and drug overdose.

Definitive etiologic diagnosis requires isolation of the microorganism from blood or a local site of infection. At least two blood samples (10 mL each) should be obtained (from different venipuncture sites) for culture. Because gram-negative bacteremia is typically low-grade (<10 organisms per milliliter of blood), multiple blood cultures or prolonged incubation of cultures may be necessary; *S. aureus* grows more readily and is detectable in blood cultures within 48 h in most instances. In many cases, blood cultures are negative; this result can reflect prior antibiotic administration, the presence of slow-growing or fastidious organisms, or the absence of microbial invasion of the bloodstream. In these cases, Gram's staining and culture of material from the primary site of infection or of infected cutaneous lesions may help establish the microbial etiology. The skin and mucosae should be examined carefully and repeatedly for lesions that might yield diagnostic information. With overwhelming bacteremia (e.g., pneumococcal sepsis in splenectomized individuals or fulminant meningococcemia), microorganisms are sometimes visible on buffy coat smears of peripheral blood.

Detection of endotoxin in blood by the limulus lysate test may portend a poor outcome, but this assay is not useful for diagnosing gram-negative bacterial infections, including gram-negative bacteremia. Although blood levels ofIL-6 also may correlate with prognosis, cytokine assays are poorly standardized and currently have limited clinical value.

TREATMENT

Patients in whom sepsis is suspected must be managed expeditiously. This task is best accomplished in an intensive care unit by personnel who are experienced in the care of the critically ill. Successful management requires urgent measures to treat the local site of infection, to provide hemodynamic and respiratory support, and to eliminate the offending microorganism. The outcome is also influenced by the patient's underlying disease, which should be managed aggressively.

Antimicrobial Agents Antimicrobial chemotherapy should be initiated as soon as samples of blood and other relevant sites have been cultured. The choice of initial therapy is based on knowledge of the likely pathogens at specific sites of local infection. Available information about patterns of antimicrobial susceptibility among bacterial isolates from the community, the hospital, and the patient also should be taken into account. It is important, pending culture results, to initiate empirical antimicrobial therapy

that is effective against both gram-positive and gram-negative bacteria (Table 124-3). Maximal recommended doses of antimicrobial drugs should be given intravenously, with adjustment for impaired renal function when necessary. When culture results become available, the regimen can often be simplified, as a single antimicrobial agent is frequently adequate for the treatment of a known pathogen. Most patients require antimicrobial therapy for at least 1 week; the duration of treatment is typically influenced by factors such as the site of tissue infection, the adequacy of surgical drainage, the patient's underlying disease, and the antimicrobial susceptibility of the bacterial isolate(s).

Removal of the Source of Infection Removal or drainage of a focal source of infection is essential. Sites of occult infection should be sought carefully. Indwelling intravenous catheters should be removed, the tip rolled over a blood agar plate for quantitative culture, and a new catheter inserted at a different site. Foley and drainage catheters should be replaced. The possibility of paranasal sinusitis (often caused by gram-negative bacteria) should be considered if the patient has undergone nasal intubation. In the neutropenic patient, cutaneous sites of tenderness and erythema, particularly in the perianal region, must be carefully sought. In patients with sacral or ischial decubitus ulcers, it is important to exclude pelvic or other soft-tissue pus collections (by computed tomography or magnetic resonance imaging, if necessary). In patients with severe sepsis arising from the urinary tract, sonography or computed tomography should be used to rule out ureteral obstruction, perinephric abscess, and renal abscess. These studies are not so urgent in patients with less severe urosepsis, provided that a clinical response is evident within 48 to 72 h.

Hemodynamic, Respiratory, and Metabolic Support (See also Chap. 38) The primary goal is to restore adequate oxygen and substrate delivery to the tissues. Adequate organ perfusion is essential. Effective intravascular volume depletion is common in patients with sepsis, and initial management of hypotension should include the administration of intravenous fluids, typically 1 to 2 L of normal saline over 1 to 2 h. The pulmonary capillary wedge pressure or the central venous pressure must be monitored in patients with refractory shock or underlying cardiac or renal disease. To avoid pulmonary edema, the pulmonary capillary wedge pressure should be maintained between 12 and 16 mmHg or the central venous pressure between 10 and 12 cmH_2O. The urine output rate should be kept above 30 mL/h by continuing fluid administration; a diuretic such as furosemide may be used if needed. In about one-third of patients, hypotension and organ hypoperfusion respond to fluid resuscitation; a reasonable goal is to maintain a mean arterial blood pressure of >60 mmHg (systolic pressure,>90 mmHg) and a cardiac index of^{3}4 (L/min)/m_2. If these guidelines cannot be met by volume infusion, inotropic and vasopressor therapy is indicated (Chap. 38). Circulatory adequacy is also assessed by clinical parameters (mentation, urine output, skin perfusion) and, when possible, by measurements of oxygen delivery and consumption.

Adrenal insufficiency should be considered in septic patients with refractory hypotension, fulminant *N. meningitidis* bacteremia, prior glucocorticoid use, disseminated tuberculosis, or AIDS. The cosyntropin (a_{1-24}-ACTH) stimulation test (Chap. 331) may suggest absolute or partial adrenal insufficiency. Supplemental hydrocortisone (50 mg intravenously every 6 h) may be given while the results of the cosyntropin test are awaited.

Ventilator therapy is indicated for progressive hypoxemia, hypercapnia, neurologic deterioration, or respiratory muscle failure. Sustained tachypnea (respiratory rate,>30 breaths/min) is frequently a harbinger of impending respiratory collapse; mechanical ventilation is often initiated to ensure adequate oxygenation, divert blood from the muscles of respiration, prevent aspiration of oropharyngeal contents, and reduce the cardiac afterload. Blood or erythrocyte transfusion is indicated if oxygen delivery is compromised by a low hemoglobin concentration (<8 to 10 g/dL).

Bicarbonate is sometimes administered for severe metabolic acidosis (arterial pH < 7.2).DIC, if complicated by major bleeding, should be treated with transfusion of fresh-frozen plasma and platelets. Successful treatment of the underlying infection is essential to reverse both acidosis and DIC.

These are consensus recommendations; none of these generally accepted components of resuscitative care has been validated in randomized clinical trials.

General Support In patients with prolonged severe sepsis (i.e., that lasting more than 2 or 3 days), nutritional supplementation may reduce the impact of protein hypercatabolism; available evidence favors the enteral delivery route. Recovery is also assisted by preventing skin breakdown, deep venous thrombosis, nosocomial infections, and stress ulcers.

Other Measures Despite aggressive management, many patients with severe sepsis or septic shock die. Two kinds of agents that may help prevent these deaths are being investigated: (1) drugs that neutralize bacterial endotoxin, thereby potentially benefiting the fraction (approximately half) of septic patients who have gram-negative bacterial infection, and (2) drugs that interfere with one or more mediators of the inflammatory response and may benefit all patients with sepsis.

Antiendotoxin Agents Lipid A, the toxic moiety of endotoxin, is conserved in theLPS of gram-negative bacteria. Despite much effort to develop drugs that bind lipid A and neutralize endotoxin in vivo, the potential of endotoxin as a target for therapeutic intervention remains controversial. In placebo-controlled clinical trials, two monoclonal antibodies to endotoxin did not prevent the death of patients with severe gram-negative bacterial sepsis. In retrospective studies, these antibodies did not bind to LPS with high affinity, and one was reported to be a polyreactive autoantibody. A theoretically more promising agent is bactericidal permeability-increasing protein, a human neutrophil protein that neutralizes the toxicity of lipid A and may be bactericidal to many gram-negative bacteria. In one clinical trial, this protein decreased morbidity and mortality among children with fulminant meningococcemia. Other investigational drugs include nontoxic lipid A analogs that reduce host responses to endotoxins and lipoproteins (such as high-density lipoprotein) that bind and neutralize endotoxin in the circulation and can remove it from the surfaces of myeloid cells.

Antimediator Agents Other adjunctive therapies are intended to control the inflammatory response, regardless of the microbial stimulus. However, numerous agents that directly or indirectly interfere with the actions of inflammatory mediators have not prevented the death of patients with severe sepsis or septic shock. Many

factors have probably contributed to the unsuccessful outcomes of these trials, including problems with study design (inappropriate end points, inadequate sample size, population heterogeneity, multiple covariates) and drug administration (wrong dose, time, or duration of administration). Anti-inflammatory drugs tested in clinical trials include methylprednisolone, ibuprofen, recombinant IL-1Ra, genetically engineered soluble receptors forTNF-a, and monoclonal antibodies to TNF-a. Because TNF-a andIL-1bdoubtless play key roles in antimicrobial host defense, neutralizing these cytokines could be detrimental in some cases. In addition, studies suggest that many anti-inflammatory molecules, including soluble TNF receptors and IL-1Ra, may already be present at high concentrations in the plasma of patients with septic shock. Identifying beneficial regimens of treatment with drugs that neutralize TNF-a and IL-1bmay therefore be very difficult. Clinical trials are testing drugs (antithrombin, activated protein C, tissue factor pathway inhibitor) intended to prevent or reverse microthombosis and evaluating regimens in which low doses of glucocorticoids are administered for prolonged periods.

All of the more recent clinical trials have enrolled patients with severe sepsis or septic shock. Neither the ability of adjunctive agents to prevent severe sepsis or septic shock in high-risk patients nor the value of combination therapy with two or more adjunctive drugs has been tested.

PROGNOSIS

Approximately 20 to 35% of patients with severe sepsis and 40 to 60% of patients with septic shock die within 30 days. Others die within the ensuing 6 months. Late deaths often result from poorly controlled infection, complications of intensive care, failure of multiple organs, or the patient's underlying disease.

Several prognostic stratification systems indicate that factoring in the patient's age, underlying condition, and various physiologic variables can yield estimates of the risk of dying of severe sepsis. Of the individual covariates, the severity of underlying disease most strongly influences the risk of dying. Septic shock is also a strong predictor of short- and long-term mortality. Case-fatality rates are similar for culture-positive and culture-negative severe sepsis.

PREVENTION

Prevention offers the best opportunity to reduce morbidity and mortality. Most episodes of severe sepsis and septic shock are nosocomial. These cases might be prevented by reducing the number of invasive procedures undertaken, by limiting the use (and duration of use) of indwelling vascular and bladder catheters, by reducing the incidence and duration of profound neutropenia (<500 neutrophils/uL), and by more aggressively treating localized nosocomial infections. Indiscriminate use of antimicrobial agents and glucocorticoids should be avoided, and optimal infection-control measures (Chap. 134) should be used. In addition, prompt and aggressive management of patients with sepsis is imperative. Studies indicate that 50 to 70% of patients who develop nosocomial severe sepsis or septic shock have experienced a less severe stage of the septic response (e.g.,SIRS, sepsis) on at least one previous day in the hospital. Research is needed to identify patients at high risk for severe sepsis and to develop adjunctive

agents that can damp the septic response before organ dysfunction or hypotension occurs.

(Bibliography omitted in Palm version)

Back to Table of Contents

125. FEVER OF UNKNOWN ORIGIN - *Jeffrey A. Gelfand*

DEFINITION AND CLASSIFICATION

Fever of unknown origin (FUO) was defined by Petersdorf and Beeson in 1961 as (1) temperatures >38.3°C (101°F) on several occasions; (2) a duration of fever of >3 weeks; and (3) failure to reach a diagnosis despite 1 week of inpatient investigation. While this classification has stood for more than 30 years, Durack and Street have proposed a new system for classification of FUO: (1) classic FUO; (2) nosocomial FUO; (3) neutropenic FUO; and (4) FUO associated with HIV infection (Table 125-1).

Classic FUO ClassicFUOcorresponds closely to the earlier definition of FUO, differing only with regard to the prior requirement for 1 week's study in the hospital. The new definition is broader, stipulating three outpatient visits or 3 days in the hospital without elucidation of a cause or 1 week of "intelligent and invasive" ambulatory investigation.

Nosocomial FUO In nosocomialFUO, a temperature of³38.3°C (101°F) develops on several occasions in a hospitalized patient who is receiving acute care and in whom infection was not manifest or incubating on admission. Three days of investigation, including at least 2 days' incubation of cultures, is the minimum requirement for this diagnosis.

Neutropenic FUO NeutropenicFUO is defined as a temperature of³38.3°C (101°F) on several occasions in a patient whose neutrophil count is <500/uL or is expected to fall to that level in 1 to 2 days. The diagnosis of neutropenic FUO is invoked if a specific cause is not identified after 3 days of investigation, including at least 2 days' incubation of cultures.

HIV-Associated FUO FUOassociated with HIV infection is defined by a temperature of³38.3°C (101°F) on several occasions over a period of>4 weeks for outpatients or>3 days for hospitalized patients with HIV infection. This diagnosis is invoked if appropriate investigation over 3 days, including 2 days' incubation of cultures, reveals no source.

Adoption of these categories ofFUO on a wide scale in the literature would allow a more rational compilation of data regarding these disparate groups. In the remainder of this chapter, the discussion will focus on classic FUO unless otherwise specified.

CAUSES OF CLASSIC FUO

Table 125-2 summarizes the findings of several large studies ofFUOcarried out since the advent of the antibiotic era, including a prospective study of 167 adult patients with FUO encompassing all 8 university hospitals in the Netherlands and using a standardized protocol in which the first author reviewed every patient. Coincident with the widespread use of antibiotics, increasingly useful diagnostic technologies -- both noninvasive and invasive -- have been developed. Newer studies reflect not only changing patterns of disease but also the impact of diagnostic techniques that make it possible to eliminate many patients with specific illness from the FUO category. The ubiquitous use of microbiologic cultures and the widespread use of potent broad-spectrum antibiotics may have decreased the number of infections causing FUO.

The wide availability of ultrasonography, computed tomography (CT), and magnetic resonance imaging (MRI) has enhanced the detection of occult neoplasms and lymphomas in patients previously thought to have FUO. Likewise, the widespread availability of highly specific and sensitive immunologic testing has reduced the number of undetected cases of systemic lupus erythematosus and other autoimmune diseases.

Several generalizations can be made. Infections, especially extrapulmonary tuberculosis, remain the leading diagnosable cause ofFUO. Prolonged mononucleosis syndromes caused by Epstein-Barr virus, cytomegalovirus (CMV), or HIV are conditions whose consideration as a cause of FUO is sometimes confounded by delayed antibody responses. Intraabdominal abscesses (sometimes poorly localized) and renal, retroperitoneal, and paraspinal abscesses continue to be difficult to diagnose. Renal malacoplakia, with submucosal plaques or nodules involving the urinary tract, may cause FUO and is often fatal if untreated. It is associated with coliform infection, is seen most often in patients with defects of intracellular bacterial killing, and is treated with fluoroquinolones or trimethoprim-sulfamethoxazole. Occasionally, other organs may be involved. Osteomyelitis, especially where prosthetic devices have been implanted, and infective endocarditis must be considered. Although true culture-negative infective endocarditis is rare, one may be misled by slow-growing organisms of the HACEK group (*Haemophilus aphrophilus*, *Actinobacillus actinomycetemcomitans*, *Cardiobacterium hominis*, *Eikenella corrodens*, and *Kingella kingae*;Chap. 150), *Bartonella* spp. (previously *Rochalimaea*), *Legionella* spp., *Coxiella burnetii*, *Chlamydia psittaci*, and fungi. Prostatitis, dental abscesses, sinusitis, and cholangitis continue to be sources of occult fever.

Fungal disease, most notably histoplasmosis involving the reticuloendothelial system, may causeFUO. FUO with headache should prompt examination of spinal fluid for *Cryptococcus neoformans*. Malaria (which may result from transfusion, the failure to take a prescribed prophylactic agent, or infection with a drug-resistant strain) continues to be a cause, particularly of nonsynchronized FUO. A related protozoan species, *Babesia*, may cause FUO and is increasing in incidence.

In most earlier series, neoplasms were the next most common cause ofFUOafter infections (Table 125-3). In the two most recent series, a decrease in the percentage of FUO cases due to malignancy was attributed to improvement in diagnostic technologies. This observation does not diminish the importance of considering neoplasia in the initial diagnostic evaluation of a patient with fever. A number of patients in these series had temporal arteritis, adult Still's disease, drug-related fever, and factitious fever. In recent series, approximately 25 to 30% of cases of FUO have remained undiagnosed. The general term *noninfectious inflammatory diseases* applies to systemic rheumatologic or vasculitic diseases such as polymyalgia rheumatica, lupus, and adult Still's disease as well as to granulomatous diseases such as sarcoidosis and Crohn's and granulomatous hepatitis.

In the elderly, multisystem disease is the most frequent cause ofFUO, giant cell arteritis being the leading etiologic entity in this category. Tuberculosis is the most common infection causing FUO in the elderly, and colon cancer is an important cause of FUO with malignancy.

Many diseases have been grouped in the various studies as "miscellaneous." On this list are drug fever, pulmonary embolism, factitious fever, familial Mediterranean fever, and Fabry's disease.

A drug-related etiology must be considered in any case of prolonged fever. Any febrile pattern may be elicited by a drug, and both relative bradycardia and hypotension are uncommon. Eosinophilia and/or rash is found in only one-fifth of patients with drug fever, which usually begins 1 to 3 weeks after the start of therapy and remits 2 to 3 days after therapy is stopped. Virtually all classes of drugs cause fever, but antimicrobials (especially b-lactam antibiotics), cardiovascular drugs (e.g., quinidine), antineoplastic drugs, and drugs acting on the central nervous system (e.g., phenytoin) are particularly common causes.

It is axiomatic that, as the duration of fever increases, the likelihood of an infectious cause decreases (Table 125-4). In a series of 347 patients referred to the National Institutes of Health from 1961 to 1977, only 6% had an infection. A significant proportion (9%) had factitious fevers -- i.e., fevers due either to false elevations of temperature or to self-induced disease. A substantial number of these factitious cases were in young women in the health professions. It is worth noting that 8% of the patients with prolonged fevers (some of whom had completely normal liver function studies) had granulomatous hepatitis, and 6% had adult Still's disease. After prolonged investigation, 19% of cases still had no specific diagnosis. A total of 27% of patients either had no actual fever during the weeks of inpatient observation or had an exaggerated circadian temperature rhythm without chills, elevated pulse, or other abnormalities.

The conditions that may be considered in a differential diagnosis of classicFUO in adults are listed in Table 125-5. This list applies strictly to the United States; the frequency of global travel underscores the need for a detailed travel history, and the continuing emergence of new infectious diseases makes this listing potentially incomplete.

SPECIALIZED DIAGNOSTIC STUDIES

Classic FUO Certain specific diagnostic maneuvers become critical in dealing with prolonged fevers. If factitious fever is suspected, electronic thermometers should be used, temperature-taking should be supervised, and simultaneous urine and body temperatures should be measured. Any tissue removed during prior relevant surgery should be reexamined; slides should be requested, and, if need be, paraffin blocks of fixed pathologic material should be reexamined and additional special studies performed. Relevant x-rays should be reexamined; reviewing of prior radiologic reports may be insufficient. Serum should be set aside in the laboratory as soon as possible and retained for future examination for rising antibody titers. *Febrile agglutinins* is a vague term that in most laboratories refers to serologic studies for salmonellosis, brucellosis, and rickettsial diseases. These studies are seldom useful, having low sensitivity and variable specificity. Rising titers of antibody to *Brucella* (Chap. 160) are usually diagnostic, but false-positive results may be obtained in typhoid fever, tularemia, and yersinial infections. Infection with *Brucella canis* may be missed with standard antibody tests for *Brucella*. *Salmonella* infection (Chap. 156) elevates antibody titers to the H and O antigens. High titers of antibody to the H antigen persist for years and may reflect previous infection or immunization. Serology for *Yersinia enterocolitica* may be

useful. The measurement of specific antirickettsial titers should be requested for the diagnosis of Rocky Mountain spotted fever and Q fever. Multiple blood samples (no fewer than three, rarely more than six), including samples for anaerobic culture, should be cultured in the laboratory for at least 2 weeks to ensure that any HACEK-group organisms that may be present have ample time to grow (Chap. 150). Lysis-centrifugation blood culture techniques should be employed in cases where prior antimicrobial therapy or fungal or atypical mycobacterial infection is suspected. Blood culture media should be supplemented with L-cysteine or pyridoxal to assist in the isolation of nutritionally variant streptococci. It should be noted that sequential cultures positive for multiple organisms may reflect self-injection of contaminated substances. Urine cultures, including cultures for mycobacteria, fungi, andCMV, are indicated. Blood, urine, or cerebrospinal fluid (CSF) can now be tested for a variety of pathogens such as CMV or hepatitis C virus by using the polymerase chain reaction (PCR) to amplify and hence detect viral nucleic acid (Chap. 121). Liver biopsy, even when the results of liver function studies are normal, should be considered and pursued if the diagnosis remains elusive. Specimens should be cultured for mycobacteria and fungi. Likewise, bone marrow biopsy (not simple aspiration) should be used to obtain specimens for histology and culture. The blood smear should be examined for *Plasmodium*, *Babesia*, *Trypanosoma*, *Leishmania*, and *Borrelia*.

In anFUOworkup, the erythrocyte sedimentation rate (ESR) should be determined. Striking elevation of the ESR and anemia of chronic disease are frequently seen in association with giant cell arteritis or polymyalgia rheumatica, common causes of FUO in patients over 50 years of age. Still's disease is also suggested by elevations of ESR, leukocytosis, and anemia and is often accompanied by arthralgias, polyserositis (pleuritis, pericarditis), lymphadenopathy, splenomegaly, and rash. Antinuclear antibody, antineutrophil cytoplasmic antibody, rheumatoid factor, and serum cryoglobulins should be measured to rule out other collagen vascular diseases and vasculitis. Another cause of an extremely high ESR may be a false-positive value attributable to a cold agglutinin with a broad thermal amplitude. The ESR test is nonspecific, yielding values that depend on certain serum proteins (most notably fibrinogen) known to interfere with the zeta-potential that keeps erythrocytes from clumping. When fibrinogen levels go up, the zeta-potential is inhibited, erythrocytes clump, and the ESR is high. A cold agglutinin, by binding to erythrocytes, can produce a false-positive agglutinin that mimics an acute-phase response; cold agglutinins may be seen in *Mycoplasma* and Epstein-Barr virus infections and in lymphomas.

With rare exceptions, the intermediate-strength purified protein derivative (PPD) skin test should be used to screen for tuberculosis in patients with classicFUO. Concurrent control tests, such as the CMI test (Connaught Labs, Swiftwater, PA), which is especially effective, should be employed. It should be kept in mind that both the PPD skin test and control tests may yield negative results in miliary tuberculosis, sarcoidosis, Hodgkin's disease, malnutrition, or AIDS. Noninvasive procedures should include an upper gastrointestinal contrast study with small-bowel follow-through and barium enema to include the terminal ileum and cecum. Chest x-rays should be repeated if new symptoms arise. In some cases, pulmonary function studies may be necessary. A diminished carbon monoxide diffusing capacity may indicate a restrictive lung disease such as sarcoidosis, even with a normal chest x-ray. In such cases, transbronchial biopsy may prove diagnostic. Flexible colonoscopy may be advisable, since colon

carcinoma is a cause of FUO and easily escapes detection by ultrasound andCT.

CTof the chest and abdomen should be performed. If a spinal or paraspinal lesion is suspected, however,MRI is preferred. MRI may be superior to CT in demonstrating intraabdominal abscesses and aortic dissection, but the relative utility of MRI and CT in the diagnosis of FUOis unknown. At present, it appears that abdominal CT, with oral and intravenous contrast, should be used unless MRI is specifically indicated. Arteriography may be useful for patients in whom systemic necrotizing vasculitis is suspected. Saccular aneurysms may be seen, most commonly in renal or hepatic vessels, and may permit diagnosis of arteritis when biopsy is difficult.Figure 125-1 shows a renal angiogram of a patient with polyarteritis nodosa. Ultrasonography of the abdomen is useful for the investigation of the hepatobiliary tract, kidneys, spleen, and pelvis. Echocardiography may be helpful in an evaluation for bacterial endocarditis, pericarditis, nonbacterial thrombotic endocarditis, and atrial myxomas. Transesophageal echocardiography is especially sensitive for these lesions.

Radionuclide scanning procedures using technetium (Tc) 99m sulfur colloid, gallium (Ga) 67 citrate, or indium (In) 111-labeled leukocytes or immunoglobulin may be useful in identifying and/or localizing inflammatory processes. In a recent study, Ga scintigraphy yielded useful diagnostic information in almost one-third of cases, and it was suggested that this procedure might actually be used before other imaging techniques if no specific organ is suspected of being abnormal. Tc bone scan should be undertaken to look for osteomyelitis or bony metastases;^{67}Ga scan may be used to identify sarcoidosis (Chap. 318) or *Pneumocystis carinii* (Chap. 209) in the lungs or Crohn's disease (Chap. 287) in the abdomen.^{111}In-labeled white blood cell (WBC) scan may be used to locate abscesses;^{111}In-labeled immunoglobulin scan also shows promise in this regard. With^{67}Ga,^{111}In-WBC, and^{111}In-immunoglobulin scans, false-positive and false-negative findings are common.

Biopsy of the liver and bone marrow should be considered routine in the workup ofFUO if the studies mentioned above are unrevealing or if fever is prolonged. It goes without saying that areas of suspected abnormality should be sampled for pathologic examination whenever practical. When possible, a section of the tissue block should be retained for further sections or stains.PCRtechnology makes it possible to identify and speciate mycobacterial DNA in paraffin-embedded, fixed tissues. Thus, in some cases, it is possible to make a retrospective diagnosis based on studies of long-fixed pathologic tissues. In a patient over age 50 (or occasionally in a younger patient) with the appropriate symptoms and laboratory findings, "blind biopsy" of one or both temporal arteries may yield a diagnosis of arteritis. If noted, tenderness or decreased pulsation should guide the selection of a site for biopsy. Lymph node biopsy may be helpful if nodes are enlarged, but inguinal nodes are often palpable and are seldom diagnostically useful.

Exploratory laparotomy has been performed when all other diagnostic procedures fail but has largely been replaced by modern imaging and guided-biopsy techniques. Laparoscopic biopsy may provide more adequate guided sampling of lymph nodes or liver.

Nosocomial FUO The primary considerations in diagnosing nosocomialFUO are the

underlying susceptibility of the patient coupled with the potential complications of hospitalization. The original surgical or procedural field is the place to begin a directed physical and laboratory examination for abscesses, hematomas, or infected foreign bodies. More than 50% of patients with nosocomial FUO are infected, and intravascular lines, septic phlebitis, and prostheses are all suspect. In this setting, the approach is to focus on sites where occult infections may be sequestered, such as the sinuses of intubated patients or a prostatic abscess in a man with a urinary catheter. *Clostridium difficile* colitis may be associated with fever and leukocytosis before the onset of diarrhea. In approximately 25% of patients with nosocomial FUO, the fever has a noninfectious cause. Among these causes are acalculous cholecystitis, deep vein thrombophlebitis, and pulmonary embolism. Drug fever, transfusion reactions, alcohol/drug withdrawal, adrenal insufficiency, thyroiditis, pancreatitis, gout, and pseudogout are among the many possible causes to consider. As in classic FUO, repeated meticulous physical examinations, coupled with focused diagnostic techniques, are imperative. Multiple blood, wound, and fluid cultures are mandatory. The pace of diagnostic tests is accelerated, and the threshold for procedures -- CTscans, ultrasonography,^{111}In-WBC scans, noninvasive venous studies -- is low. Even so, 20% of cases of nosocomial FUO may go undiagnosed.

Like diagnostic measures, therapeutic maneuvers must be swift and decisive, as many patients are already critically ill. Intravenous lines must be changed (and cultured), drugs stopped for 72 h, and empirical therapy started if bacteremia is a threat. In many hospital settings, empirical antibiotic coverage for nosocomial FUO now includes vancomycin for methicillin-resistant *Staphylococcus aureus* as well as broad-spectrum gram-negative coverage with piperacillin/tazobactam, ticarcillin/clavulanate, imipenem, or meropenem. Practice guidelines covering many of these issues have been published jointly by the Infectious Diseases Society of America (IDSA) and the Society for Critical Care Medicine and can be accessed on the IDSA website (http://www.idsociety.org/practice/index.html).

Neutropenic FUO (See also Chap. 85) Neutropenic patients are susceptible to focal bacterial and fungal infections, to bacteremic infections, to infections involving catheters (including septic thrombophlebitis), and to perianal infections. *Candida* and *Aspergillus* infections are common. Infections due to herpes simplex virus orCMV are sometimes causes ofFUO in this group. While the duration of illness may be short in these patients, the consequences of untreated infection may be catastrophic, with 50 to 60% infected, and 20% bacteremic. TheIDSA has published extensive practice guidelines covering these critically ill neutropenic patients; these guidelines appear on the website cited in the previous section. In these patients, severe mucositis, quinolone prophylaxis, colonization with methicillin-resistant *S. aureus*, obvious catheter-related infection, or hypotension would dictate the use of vancomycin plus ceftazidime or imipenem to provide empirical coverage for bacterial sepsis.

HIV-Associated FUO HIV infection alone may be a cause of fever. Infection due to *Mycobacterium avium* or *Mycobacterium intracellulare*, tuberculosis, toxoplasmosis,CMVinfection, *P. carinii* infection, salmonellosis, cryptococcosis, histoplasmosis, non-Hodgkin's lymphoma, and (of particular importance) drug fever are all possible causes ofFUO. Mycobacterial infection can be diagnosed by blood cultures and by liver, bone marrow, and lymph node biopsies. ChestCTshould be performed to

identify enlarged mediastinal nodes. Serologic studies may reveal cryptococcal antigen, and ^{67}Ga scan may help identify *P. carinii* pulmonary infection. More than 80% of HIV patients with FUO are infected, but drug fever and lymphoma remain important considerations.**Treatment of HIV-associated FUO depends on many factors and is discussed in Chap. 309.*

TREATMENT

The emphasis in patients with classicFUO is on continued observation and examination, with the avoidance of "shotgun" empirical therapy. Empirical treatment for endocarditis, for example, should be avoided unless there are specific reasons beyond fever to invoke this diagnosis. Every patient with FUO should undergo an exhaustive examination for tuberculosis. If thePPD skin test is positive or if granulomatous hepatitis or other granulomatous disease is present with anergy (and sarcoid seems unlikely), then a therapeutic trial with isoniazid and rifampin (and possibly a third drug) should be undertaken, with treatment usually continued for up to 6 weeks. A failure of the fever to respond over this period suggests an alternative diagnosis.

The response of rheumatic fever and Still's disease to aspirin and nonsteroidal anti-inflammatory agents (NSAIDs) may be dramatic. The effects of glucocorticoids on temporal arteritis, polymyalgia rheumatica, and granulomatous hepatitis are equally dramatic. Colchicine is highly effective in preventing attacks of familial Mediterranean fever but is of little use once an attack is well under way. The ability of glucocorticoids and NSAIDs to mask fever while permitting the spread of infection dictates that their use be avoided unless infection has been largely ruled out and unless inflammatory disease is both probable and debilitating or threatening.

When no underlying source ofFUO is identified after prolonged observation (>6 months), the prognosis is generally good, however vexing the fever may be to the patient. Under such circumstances, debilitating symptoms are treated withNSAIDs, and glucocorticoids are the last resort. The initiation of empirical therapy does not mark the end of the diagnostic workup; rather, it commits the physician to continued thoughtful reexamination and evaluation. Patience, compassion, equanimity, and intellectual flexibility are indispensable attributes for the clinician in dealing successfully with FUO.

ACKNOWLEDGEMENT

Sheldon M. Wolff, MD, now deceased, was an author of a previous version of this chapter. It is to his memory that the chapter is dedicated.

(Bibliography omitted in Palm version)

Back to Table of Contents

126. INFECTIVE ENDOCARDITIS - *Adolf W. Karchmer*

INTRODUCTION

The proliferation of microorganisms on the endothelium of the heart results in infective endocarditis. The prototypic lesion at the site of infection, the *vegetation* (see Plate IID-59,Fig. 126-CD1), is a mass of platelets, fibrin, microcolonies of microorganisms, and scant inflammatory cells. Infection most commonly involves heart valves (either native or prosthetic) but may also occur on the low-pressure side of the ventricular septum at the site of a defect, on the mural endocardium where it is damaged by aberrant jets of blood or foreign bodies, or on intracardiac devices themselves. The analogous process involving arteriovenous shunts, arterioarterial shunts (patent ductus arteriosus), or a coarctation of the aorta is called *infective endarteritis*.

Endocarditis may be classified according to the temporal evolution of disease, the site of infection, the cause of infection, or a predisposing risk factor such as injection drug use. While each classification criterion provides therapeutic and prognostic insight, the methods overlap and none is sufficient alone. The classification of endocarditis as acute and subacute was initially used to describe the illness and the time elapsed until death; presently it is applied to the features and progression of infection until diagnosis. *Acute endocarditis* is a hectically febrile illness, rapidly damages cardiac structures, hematogenously seeds extracardiac sites, and, if untreated, progresses to death within weeks. *Subacute endocarditis* follows an indolent course; causes structural cardiac damage only slowly, if at all; rarely causes metastatic infection; and is gradually progressive unless complicated by a major embolic event or ruptured mycotic aneurysm.

In developed countries, the incidence of endocarditis ranges from 1.5 to 6.2 cases per 100,000 population per year. In the late 1980s in a metropolitan area of the United States (Philadelphia), endocarditis occurred in 9.3 persons per 100,000 population per year. However, half of these cases arose as a consequence of injection drug use. The incidence of endocarditis is notably increased among the elderly. The cumulative rate of prosthetic valve endocarditis is 1.5 to 3.0% at 1 year after valve replacement and 3 to 6% at 5 years; the risk is greatest during the first 6 months after valve replacement.

ETIOLOGY

A vast array of microorganisms, including many species of bacteria and fungi, have been reported to cause sporadic episodes of endocarditis. Nevertheless, a small number of bacterial species cause the majority of cases (Table 126-1). The causative microorganisms vary somewhat among the major clinical types of endocarditis, in part because of the different portals of entry. The oral cavity, skin, and upper respiratory tract are the respective primary portals for the viridans streptococci, staphylococci, and HACEK organisms (*Haemophilus*, *Actinobacillus*, *Cardiobacterium*, *Eikenella*, and *Kingella*) causing community-acquired native valve endocarditis. *Streptococcus bovis* originates from the gastrointestinal tract, where it is associated with polyps and colonic tumors, and enterococci enter the bloodstream from the genitourinary tract. Nosocomial native valve endocarditis is largely the consequence of bacteremia arising from intravascular catheters and less commonly from nosocomial wound and urinary tract

infection. Endocarditis complicates 6 to 25% of episodes of catheter-associated *Staphylococcus aureus* bacteremia; higher rates are detected by careful transesophageal echocardiography (TEE) screening (see "Echocardiography," below).

Prosthetic valve endocarditis arising within 2 months of valve surgery is generally the result of intraoperative contamination of the prosthesis or a bacteremic postoperative complication. The nosocomial nature of these infections is reflected in their primary microbial causes: coagulase-negative staphylococci, *S. aureus*, facultative gram-negative bacilli, diphtheroids, and fungi. The portals of entry and organisms causing cases beginning >12 months after surgery are similar to those in community-acquired native valve endocarditis. Epidemiologic evidence suggests that prosthetic valve endocarditis due to coagulase-negative staphylococci that presents between 2 and 12 months after surgery is often nosocomial in origin but with a delayed onset. At least 85% of coagulase-negative staphylococci that cause prosthetic valve endocarditis within 12 months of surgery are methicillin-resistant; the rate of methicillin resistance decreases to 25% among coagulase-negative staphylococci causing prosthetic endocarditis that presents >1 year after valve surgery.

Transvenous pacemaker lead and/or implanted defibrillator-associated endocarditis is usually a nosocomial infection. The majority of episodes occur within weeks of implantation or generator change and are caused by *S. aureus* or coagulase-negative staphylococci.

Endocarditis occurring among injection drug users, especially when infection involves the tricuspid valve, is commonly caused by *S. aureus* strains, many of which are methicillin-resistant. The causes of left-sided valve infection in addicts are more varied, and the involved valves have often been damaged by prior episodes of endocarditis. A number of these cases are caused by *Pseudomonas aeruginosa* and *Candida* species, and sporadic cases are due to unusual organisms such as *Bacillus*, *Lactobacillus*, and *Corynebacterium* species. Polymicrobial endocarditis occurs more frequently in injection drug users than in patients who do not inject drugs. The presence of HIV in this population does not significantly impact the causes of endocarditis.

From 5 to 15% of patients with endocarditis have negative blood cultures; in one-third to one-half of these cases, cultures are negative because of prior antibiotic exposure. The remainder of these patients are infected by fastidious organisms, such as pyridoxal-requiring streptococci (now designated *Abiotrophia* species), the gram-negative coccobacillary HACEK organisms, *Bartonella henselae*, or *Bartonella quintana*. Some fastidious organisms that cause endocarditis have characteristic epidemiologic settings (e.g., *Coxiella burnetii* in Europe, *Brucella* species in the Middle East). *Tropheryma whippelii* causes an indolent, culture-negative, afebrile form of endocarditis.

PATHOGENESIS

Unless it is injured, the normal endothelium is resistant to infection by most bacteria and to thrombus formation. Endothelial injury (e.g., at the site of impact of high-velocity jets or on the low-pressure side of a cardiac structural lesion) causes aberrant flow and allows either direct infection by virulent organisms or the development of an uninfected

platelet-fibrin thrombus -- a condition called *nonbacterial thrombotic endocarditis* (NBTE). The thrombus subsequently serves as a site of bacterial attachment during transient bacteremia. The cardiac lesions most commonly resulting in NBTE are mitral regurgitation, aortic stenosis, aortic regurgitation, ventricular septal defects, and complex congenital heart disease. These lesions result from rheumatic heart disease (particularly in the developing world, where rheumatic fever remains prevalent), mitral valve prolapse, degenerative heart disease, and congenital malformations. NBTE also arises as a result of a hypercoagulable state; this phenomenon gives rise to the clinical entity of *marantic endocarditis* (seen in patients with malignancy) and to bland vegetations complicating systemic lupus erythematosus and the antiphospholipid antibody syndrome.

Organisms that cause endocarditis generally enter the bloodstream from mucosal surfaces, the skin, or sites of focal infection. Except for more virulent bacteria (e.g., *S. aureus*) that can adhere directly to intact endothelium or exposed subendothelial tissue, microorganisms in the blood adhere to thrombi. If resistant to the bactericidal activity of serum and the microbicidal peptides released by platelets, the organisms proliferate and either stimulate tissues to produce a procoagulant or themselves exert procoagulant activity leading to further platelet-fibrin deposition and vegetation formation. Although an enormous variety of microorganisms circulate transiently in the bloodstream, only a limited number commonly cause endocarditis. The etiologic organisms of endocarditis bear surface components that facilitate adherence to injured endothelium and host proteins or to thrombi. Experiments suggest that fibronectin receptors present on many gram-positive bacteria, clumping factor (a fibrinogen-binding surface protein) on *S. aureus*, and dextrans on streptococci facilitate adherence. Organisms become enmeshed in the growing platelet-fibrin vegetation and, in the absence of host defenses, proliferate to form dense microcolonies. More than 90% of the organisms in vegetations are metabolically inactive (nongrowing) and thus are relatively resistant to killing by antimicrobial agents. Proliferating surface organisms are shed into the bloodstream continuously, whereupon some are cleared by the reticuloendothelial system and others are redeposited on the vegetation and stimulate further vegetation growth.

The pathophysiologic consequences and clinical manifestations of endocarditis -- other than constitutional symptoms, which are probably a result of cytokine production -- arise from damage to intracardiac structures; embolization of vegetation fragments, leading to infection or infarction of remote tissues; hematogenous infection of sites during bacteremia; and tissue injury due to the deposition of circulating immune complexes or immune responses to deposited bacterial antigens.

CLINICAL MANIFESTATIONS

The clinical syndrome of infective endocarditis is highly variable, may involve multiple organs, and spans a continuum between acute and subacute presentations. Native valve endocarditis (whether acquired in the community or nosocomially), prosthetic valve endocarditis, and endocarditis due to injection drug use share clinical and laboratory manifestations (Table 126-2). Although the relationship is not absolute, the causative microorganism is primarily responsible for the temporal course of endocarditis. b-Hemolytic streptococci, *S. aureus*, and pneumococci typically result in an acute course, although *S. aureus* occasionally causes subacute disease.

Endocarditis caused by *Staphylococcus lugdunensis* (a coagulase-negative species) or by enterococci may present acutely. Subacute endocarditis is typically caused by viridans streptococci, enterococci, coagulase-negative staphylococci, and the HACEK group. Endocarditis caused by *Bartonella* species and the agent of Q fever, *C. burnetii*, is exceptionally indolent.

The clinical features of endocarditis are nonspecific. However, these symptoms in a febrile patient with valvular abnormalities or a behavior pattern (injection drug use) that predisposes to endocarditis suggest the diagnosis, as do bacteremia with organisms that frequently cause endocarditis, otherwise unexplained arterial emboli, and progressive cardiac valvular incompetence. In patients with subacute presentations, fever is typically low-grade and rarely exceeds 39.4°C (103°F); in contrast, temperatures between 39.4 and 40°C (103 and 104°F) are often noted in acute endocarditis. Fever may be blunted or absent in patients who are elderly or severely debilitated or who have marked cardiac or renal failure.

Cardiac Manifestations Although heart murmurs are usually indicative of the predisposing cardiac pathology rather than of endocarditis, valvular damage and ruptured chordae may result in new regurgitant murmurs. In acute endocarditis involving a normal valve, murmurs are heard on presentation in only 30 to 45% of patients but ultimately are detected in 85%. Congestive heart failure develops in 30 to 40% of patients; it is usually a consequence of valvular dysfunction but occasionally is due to endocarditis-associated myocarditis or an intracardiac fistula. The temporal progression of heart failure is variable and depends upon the severity of valvular dysfunction; failure due to aortic valve dysfunction progresses more rapidly than that due to mitral valve dysfunction. Extension of infection beyond valve leaflets into adjacent annular or myocardial tissue results in perivalvular abscesses, which in turn may cause fistulae (from the root of the aorta into cardiac chambers or between cardiac chambers) with new murmurs. Abscesses may burrow from the aortic valve annulus through the epicardium, causing pericarditis. Extension of infection into paravalvular tissue adjacent to either the right or the noncoronary cusp of the aortic valve may interrupt the conduction system in the upper interventricular septum, leading to varying degrees of heart block. Although perivalvular abscesses arising from the mitral valve may potentially interrupt conduction pathways near the atrioventricular node or in the proximal bundle of His, such interruption occurs infrequently. Emboli to a coronary artery may result in myocardial infarction; nevertheless, embolic transmural infarcts are rare.

Noncardiac Manifestations (Figs. 19-CD2,124-CD2, and126-CD2) The classic nonsuppurative peripheral manifestations of subacute endocarditis are related to the duration of infection and, with early diagnosis and treatment, have become infrequent. In contrast, septic embolization mimicking some of these lesions (subungual hemorrhage, (Fig. 19-CD4). Osler's nodes,Fig. 126-CD3) is common in patients with acute *S. aureus* endocarditis (see Plate IID-58). Musculoskeletal symptoms, including nonspecific inflammatory arthritis and back pain, usually remit promptly with treatment but must be distinguished from the symptoms of focal metastatic infection. Hematogenously seeded focal infection may involve any organ but most often is clinically evident in the skin, spleen, kidneys, skeletal system, and meninges. Arterial emboli, which may be asymptomatic and discovered only at autopsy, are clinically

apparent in up to 50% of patients. Vegetations>10 mm in diameter (as measured by echocardiography) and those located on the mitral valve are more likely to embolize than are smaller or nonmitral vegetations. Embolic events -- often with infarction -- involving the extremities, spleen, kidneys (Fig. 126-1), bowel, or brain are often noted at presentation. With antibiotic treatment, the frequency of embolic events decreases from 13 per 1000 patient-days during the initial week to 1.2 per 1000 patient-days after the third week. Emboli occurring late during or after effective therapy do not in themselves constitute evidence of failed antimicrobial treatment. Neurologic symptoms, most often resulting from embolic strokes, occur in up to 40% of patients. Other neurologic complications include aseptic or purulent meningitis, intracranial hemorrhage due to hemorrhagic infarcts or ruptured mycotic aneurysms, seizures, and encephalopathy. Microabscesses in brain and meninges occur commonly in *S. aureus* endocarditis; surgically drainable abscesses are infrequent.

Immune complex deposition on the glomerular basement membrane causes diffuse hypocomplementemic glomerulonephritis and renal dysfunction, which typically improve with effective antimicrobial therapy. Embolic renal infarcts cause flank pain and hematuria but rarely cause renal dysfunction.

Manifestations of Specific Predisposing Conditions Among injection drug users, infection involving valves on the left side of the heart presents with the typical clinical features of endocarditis. In almost 50% of patients with endocarditis associated with injection drug use, infection is limited to the tricuspid valve. These patients present with fever, faint or no murmur, and (in 75% of cases) prominent pulmonary findings, including cough, pleuritic chest pain, nodular pulmonary infiltrates, and occasionally pyopneumothorax.

Nosocomial endocarditis (defined as that which results from hospital care within the prior month and most commonly presenting as intravascular catheter-associated bacteremia), if not associated with a retained intracardiac device, has typical manifestations. Endocarditis associated with flow-directed pulmonary artery catheters is often cryptic, with symptoms masked by comorbid critical illness, and is commonly diagnosed at autopsy. Transvenous pacemaker lead and/or implanted defibrillator-associated endocarditis commonly follows initial implantation or a generator unit change; may be associated with obvious or cryptic generator pocket infection; and results in fever, minimal murmur, and pulmonary symptoms similar to those encountered in addicts with tricuspid endocarditis.

Prosthetic valve endocarditis presents with typical clinical features. Cases arising within 60 days of valve surgery (early onset) lack peripheral vascular manifestations and may be obscured by comorbidity associated with recent surgery. In both early-onset and more delayed presentations, paravalvular infection is common and often results in partial valve dehiscence, regurgitant murmurs, congestive heart failure, or disruption of the conduction system.

DIAGNOSIS

The Duke Criteria The diagnosis of infective endocarditis is established with certainty only when vegetations obtained at cardiac surgery, at autopsy, or from an artery (an

embolus) are examined histologically and microbiologically. Nevertheless, a highly sensitive and specific diagnostic schema -- known as the *Duke criteria* -- has been developed on the basis of clinical, laboratory, and echocardiographic findings (Table 126-3). Documentation of two major criteria, of one major and three minor criteria, or of five minor criteria allows a clinical diagnosis of definite endocarditis. The diagnosis of endocarditis is rejected if an alternative diagnosis is established, if symptoms resolve and do not recur with £4 days of antibiotic therapy, or if surgery or autopsy after £4 days of antimicrobial therapy yields no histologic evidence of endocarditis. Illnesses not classified as definite endocarditis or rejected are considered cases of possible infective endocarditis. When pathologically confirmed cases have been scored retrospectively by these criteria, 90% fulfill the definition of definite or possible endocarditis; 10% are rejected (primarily because of an incomplete echocardiographic evaluation). In comparison with expert opinion, the Duke criteria identify cases considered to be endocarditis but also accept a small percentage of cases rejected by the experts. This potential for a false-positive diagnosis is the major deficiency in this schema when used clinically. If all patients with a diagnosis of definite or possible endocarditis are fully treated for endocarditis, this reduced specificity results in excess treatment for some patients. A modification of the schema has been proposed in order to increase its specificity without significantly reducing its sensitivity. This modification would require documentation of at least one major or three minor criteria for cases to be categorized as possible endocarditis.

The roles of bacteremia and echocardiographic findings in the diagnosis of endocarditis are appropriately emphasized in the Duke criteria. That multiple blood cultures obtained over time are positive is consistent with the known continuous low-density nature of bacteremia in patients with endocarditis (£100 organisms per milliliter). Among untreated endocarditis patients who ultimately have a positive blood culture, 95% of all blood cultures are positive, and in 98% of cases one of the initial two sets of cultures yields the microorganism. The diagnostic criteria attach significance to the species of organism isolated from blood cultures. To fulfill a major criterion, the isolation of an organism that causes both endocarditis and bacteremia in the absence of endocarditis (e.g., *S. aureus*, enterococci) must take place repeatedly (i.e., persistent bacteremia) and in the absence of a primary focus of infection. Organisms that rarely cause endocarditis but commonly contaminate blood cultures (e.g., diphtheroids, coagulase-negative species) must be isolated repeatedly if their isolation is to serve as a major criterion.

Blood Cultures Isolation of the causative microorganism from blood cultures is critical not only for diagnosis but also for determination of antimicrobial susceptibility and planning of treatment. In the absence of prior antibiotic therapy, a total of three blood culture sets, ideally with the first separated from the last by at least 1 h, should be obtained from different venipuncture sites over 24 h. If the cultures remain negative after 48 to 72 h, two or three additional blood cultures, including a lysis-centrifugation culture, should be obtained, and the laboratory should be asked to pursue fastidious microorganisms by prolonging incubation time and performing special subcultures. Empirical antimicrobial therapy should not be administered initially to hemodynamically stable patients with subacute endocarditis, especially those who have received antibiotics within the preceding 2 weeks; thus, if necessary, additional blood cultures can be obtained without the confounding effect of empirical treatment. Patients with

acute endocarditis or with deteriorating hemodynamics that may require urgent surgery should be treated empirically immediately after obtaining the initial three sets of blood cultures.

Non-Blood-Culture Tests for the Etiologic Agent Serologic tests can be used to identify some organisms causing endocarditis that are difficult to recover by blood culture: *Brucella*, *Bartonella*, *Legionella*, and *C. burnetii*. Pathogens can also be identified in vegetations by culture, by microscopic examination with special stains, and by use of polymerase chain reaction to recover unique microbial DNA or 16S rRNA.

Echocardiography Cardiac imaging with echocardiography allows anatomic confirmation of infective endocarditis, sizing of vegetations, detection of intracardiac complications, and assessment of cardiac function. A two-dimensional study with color flow and continuous as well as pulsed Doppler is optimal. Transthoracic echocardiography (TTE) is noninvasive and exceptionally specific; however, it cannot image vegetations<2 mm in diameter, and in 20% of patients it is technically inadequate because of emphysema or body habitus. Thus, TTE detects vegetations in only 65% of patients with definite clinical endocarditis (i.e., it has a sensitivity of 65%). Moreover, TTE is not adequate for evaluating prosthetic valves or detecting intracardiac complications. Transesophageal echocardiography (TEE) is safe and significantly more sensitive than TTE. It detects vegetations in>90% of patients with definite endocarditis; nevertheless, false-negative studies are noted in 6 to 18% of endocarditis patients. TEE is the optimal method for the diagnosis of prosthetic endocarditis or the detection of myocardial abscess, valve perforation, or intracardiac fistulae.

Experts favor echocardiographic evaluation of all patients with a clinical diagnosis of endocarditis; however, the test should not be used to screen patients with otherwise explained positive blood cultures or patients with unexplained fever. In patients with a low pretest likelihood of endocarditis (<5%), a high-qualityTTE that is negative is sufficient to exclude endocarditis. For patients whose habitus makes them difficult to study with TTE and for those who may have prosthetic valve endocarditis or who are at high risk of intracardiac complications,TEE is the preferred imaging modality. For patients with a pretest probability of endocarditis ranging from 5 to 50%, initial evaluation by TEE -- in lieu of a sequential strategy of TTE, which, if negative, will be followed by TEE -- is cost-effective. A negative TEE when endocarditis is likely does not exclude the diagnosis but rather warrants repetition of the study in 7 to 10 days with optimal multiplanar technique.

Other Studies Many laboratory studies that do not aid in diagnostic evaluation are nevertheless important in the management of patients with endocarditis; these studies include complete blood counts, creatinine measurement, chest radiography, and electrocardiography. The erythrocyte sedimentation rate, C-reactive protein level, circulating immune complex titer, and rheumatoid factor concentration are commonly increased in endocarditis (Table 126-2). Cardiac catheterization is useful only to assess coronary artery patency in older individuals who are to undergo surgery for endocarditis.

TREATMENT

Antimicrobial Therapy It is difficult to eradicate bacteria from the avascular vegetation

in infective endocarditis because this site is relatively inaccessible to host defenses and because the bacteria are nongrowing and metabolically inactive. Since all bacteria in the vegetation must be killed, therapy for endocarditis must be bactericidal and must be given for prolonged periods. Antibiotics are generally given parenterally and must reach high serum concentrations that will, through passive diffusion, lead to effective concentrations in the depths of the vegetation. The choice of effective therapy requires precise knowledge of the susceptibility of the causative microorganisms. The initiation of treatment before a cause is defined must balance the need to establish a microbiologic diagnosis against the potential progression of disease or the need for urgent surgery (see "Blood Cultures" above). The individual vulnerabilities of the patient should be weighed in the selection of therapy -- e.g., allergies, end-organ dysfunction, interactions with concomitant medications, and risks of adverse events.

Although given for several weeks longer, the regimens recommended for the treatment of endocarditis involving prosthetic valves (except for staphylococcal infections) are similar to those used to treat native valve infection (Table 126-4). Recommended doses and duration of therapy should be adhered to unless alterations are required by adverse events.

Organism-Specific Therapies

STREPTOCOCCI Although most strains of viridans streptococci and *S. bovis* that cause endocarditis are susceptible to penicillin [minimum inhibitory concentration (MIC) £0.1 ug/mL], recent reports indicate increasing penicillin resistance among viridans streptococci recovered from blood cultures. In the selection of optimal therapy, the penicillin MIC must be determined (Table 126-4). The 2-week penicillin/gentamicin regimen should not be used to treat complicated native valve infection or prosthetic valve endocarditis. Although small studies have suggested that a 2-week regimen of single daily doses of ceftriaxone (2 g IV) plus gentamicin (3 mg/kg) or netilmicin (4 mg/kg) is effective for penicillin-susceptible streptococcal endocarditis, the data are not sufficient to support routine use of this regimen. Penicillin/gentamicin is recommended for the treatment of endocarditis caused by group B streptococci.

ENTEROCOCCI Enterococci are resistant to oxacillin, nafcillin, and the cephalosporins and are inhibited only by penicillin, ampicillin, teicoplanin (not available in the United States), and vancomycin. To kill enterococci requires the synergistic interaction of a cell wall-active antibiotic (penicillin, ampicillin, vancomycin, or teicoplanin) that is effective at achievable serum concentrations and an aminoglycoside (gentamicin or streptomycin) to which the isolate does not exhibit high-level resistance. An isolate's resistance to cell wall-active agents or ability to replicate in the presence of gentamicin at³500 ug/mL or streptomycin at 2000 ug/mL -- a phenomenon called *high-level aminoglycoside resistance* -- indicates that the ineffective antimicrobial cannot participate in the interaction to produce killing. High-level resistance to gentamicin predicts that tobramycin, netilmicin, amikacin, and kanamycin will also be ineffective. In fact, even when enterococci are not highly resistant to gentamicin, it is difficult to predict the ability of these other aminoglycosides to participate in synergistic killing; consequently, they should not in general be used to treat enterococcal endocarditis.

Clearly, enterococci causing endocarditis must be tested for high-level resistance to

streptomycin and gentamicin, b-lactamase production, and susceptibility to penicillin and ampicillin (MIC,£16 ug/mL) and to vancomycin (MIC,£8 ug/mL). If the isolate producesb-lactamase, ampicillin/sulbactam or vancomycin can be used as the cell wall-active component; if the penicillin/ampicillin MIC is>16 ug/mL, vancomycin can be considered; and if the vancomycin MIC is >8 ug/mL, penicillin or ampicillin may be considered. Based on the absence of high-level resistance, gentamicin or streptomycin should be used as the aminoglycoside. If there is high-level resistance to both these drugs, no aminoglycoside should be given; instead, an 8- to 12-week course of a single cell wall-active agent is suggested. If single-drug therapy fails or the isolate is resistant to all of the commonly used agents, surgical treatment is advised. The role of newer agents potentially active against multidrug-resistant enterococci (quinupristin/dalfopristin, linezolid, and daptomycin) in the treatment of endocarditis has not been established.

STAPHYLOCOCCI The regimens used to treat staphylococcal endocarditis are not based upon coagulase production but rather upon the presence or absence of a prosthetic valve or foreign device, the native valve(s) involved, and the resistance of the isolate to penicillin and methicillin. Penicillinase is produced by 95% of staphylococci; thus, all isolates should be considered penicillin-resistant until shown not to produce this enzyme. The addition of gentamicin (if the isolate is susceptible) to a b-lactam antibiotic to enhance therapy for native mitral or aortic valve endocarditis is optional. Its addition hastens eradication of bacteremia but does not improve survival rates. If added, gentamicin should be limited to the initial 3 to 5 days of therapy to avoid nephrotoxicity. Gentamicin generally is not added to the vancomycin regimen in this setting.

Methicillin-susceptible *S. aureus* endocarditis that is uncomplicated and limited to the tricuspid or pulmonic valve -- a condition occurring almost exclusively in injection drug users -- can often be treated with a 2-week course that combines oxacillin or nafcillin (but not vancomycin) with gentamicin. Prolonged fevers (³5 days) during therapy suggest that these patients should receive standard therapy.

Staphylococcal prosthetic valve endocarditis is treated for 6 to 8 weeks with a multidrug regimen. Rifampin is an essential component because it kills staphylococci that are adherent to foreign material. Two other agents (selected on the basis of susceptibility testing) are combined with rifampin to prevent in vivo emergence of resistance. Because many staphylococci, particularly methicillin-resistant *S. aureus* and *S. epidermidis*, are resistant to gentamicin, the utility of gentamicin should be established before rifampin treatment is begun. If the isolate is resistant to gentamicin, another aminoglycoside or a fluoroquinolone (chosen in light of susceptibility results) should be substituted.

OTHER ORGANISMS Although penicillin is the therapy of choice for endocarditis caused by *S. pneumoniae*, therapy should be initiated with ceftriaxone and vancomycin until susceptibility to penicillin is established. *P. aeruginosa* endocarditis is treated with an antipseudomonal penicillin (ticarcillin or piperacillin) and high doses of tobramycin (8 mg/kg per day in three divided doses). Endocarditis caused by Enterobacteriaceae is treated with a potentb-lactam antibiotic plus an aminoglycoside. Corynebacterial endocarditis is treated with penicillin plus an aminoglycoside (if the organism is susceptible to the aminoglycoside) or with vancomycin, which is highly bactericidal for most strains. Therapy for *Candida* endocarditis consists of amphotericin B plus

flucytosine and early surgery; long-term (if not indefinite) suppression with fluconazole is used increasingly.

Empirical Therapy In designing and executing therapy without culture data (i.e., before culture results are known or when cultures are negative), clinical and epidemiologic clues to etiology must be weighed, and both the pathogens associated with the specific endocarditis syndrome and the hazards of suboptimal therapy must be considered. Thus, empirical therapy for acute endocarditis in an injection drug user should cover methicillin-resistant *S. aureus* and gram-negative bacilli. The initiation of treatment with vancomycin plus gentamicin immediately after blood is obtained for cultures covers these as well as many other potential causes. In treating culture-negative episodes, marantic endocarditis must be excluded and fastidious organisms sought serologically. In the absence of confounding prior antibiotic therapy, it is unlikely that *S. aureus* or enterococcal infection will present with negative blood cultures. Thus, in this situation, these organisms are not the determinants of therapy for subacute endocarditis. Blood culture-negative native valve endocarditis is treated with ceftriaxone (or ampicillin) plus gentamicin; these two antimicrobials plus vancomycin should be used if prosthetic valves are involved.

Outpatient Antimicrobial Therapy Fully compliant patients who have sterile blood cultures, are afebrile during therapy, and have no clinical or echocardiographic findings that suggest an impending complication may complete therapy as outpatients. Careful follow-up and a stable home setting are necessary, as are predictable intravenous access and selection of antimicrobials that are stable in solution.

Monitoring Antimicrobial Therapy The serum bactericidal titer -- the highest dilution of the patient's serum during therapy that kills 99.9% of the standard inoculum of the infecting organism -- is no longer recommended for assessment of patients receiving standard regimens. However, in the treatment of endocarditis caused by unusual organisms, this measurement, although not standardized and difficult to interpret, may provide a patient-specific assessment of in vivo antibiotic effect. Serum concentrations of aminoglycosides and vancomycin should be monitored.

Antibiotic toxicities, including allergic reactions, occur in 25 to 40% of patients and commonly arise during the third week of therapy. Blood tests to detect antibiotic-specific potential end-organ toxicity should be performed periodically.

In most patients effective antibiotic therapy results in subjective improvement and resolution of fever within 5 to 7 days. Blood cultures should be repeated daily until sterile, rechecked if there is recrudescent fever, and performed again 4 to 6 weeks after therapy to document cure. Blood cultures become sterile within 2 days after the start of appropriate therapy when infection is caused by viridans streptococci, enterococci, or HACEK organisms. In *S. aureus* endocarditis,b-lactam therapy results in sterile cultures in 3 to 5 days, whereas positive cultures may persist for 7 to 9 days with vancomycin treatment. When fever persists for 7 days in spite of appropriate antibiotic therapy, patients should be evaluated for paravalvular abscess and for extracardiac abscesses (spleen, kidney) or complications (embolic events). Recrudescent fever raises the question of these complications but also of drug reactions or complications of hospitalization. Serologic abnormalities (e.g., erythrocyte sedimentation rate,

rheumatoid factor) resolve slowly and do not reflect response to treatment. Vegetations become smaller with effective therapy, but at 3 months after cure half are unchanged and 25% are slightly larger.

Surgical Treatment Intracardiac and central nervous system complications of endocarditis are important causes of the morbidity and mortality associated with this infection. In some cases, effective treatment for these complications requires surgery. Most of the clinical indications for surgical treatment of endocarditis are not absolute (Table 126-5). The risks and benefits as well as the timing of surgical treatment must therefore be individualized.

Intracardiac Surgical Indications Most surgical interventions are clearly warranted by intracardiac findings, often detected by echocardiography. Because of the highly invasive nature of prosthetic valve endocarditis, as many as 40% of affected patients merit surgical treatment. In many patients, coincident rather than single intracardiac events necessitate surgery.

CONGESTIVE HEART FAILURE Moderate to severe refractory congestive heart failure caused by new or worsening valve dysfunction is the major indication for cardiac surgical treatment of endocarditis. Of patients with moderate to severe heart failure due to valve dysfunction who are treated medically, 60 to 90% die within 6 months. In the setting of similar hemodynamic dysfunction, surgical treatment is associated with mortality rates of 20 to 40% with native valve endocarditis and 35 to 55% with prosthetic valve infection. Surgery may be required to relieve functional stenosis due to large vegetations or to restore competence to damaged regurgitant valves.

PERIVALVULAR INFECTION This complication, which occurs in 10 to 15% of native valve and 45 to 60% of prosthetic valve infections, is suggested by persistent unexplained fever during appropriate therapy, new electrocardiographic conduction disturbances, and pericarditis. Extension can occur from any valve but is most common with aortic valve infection.TEE with color Doppler is the test of choice to detect perivalvular abscesses (sensitivity, ³85%). Although occasional perivalvular infections are cured medically, surgery is warranted when fever persists, fistulae develop, prostheses are dehisced and unstable, and infection relapses after appropriate treatment. Cardiac rhythm must be monitored since high-grade heart block may require insertion of a pacemaker.

UNCONTROLLED INFECTION Continued positive blood cultures or otherwise unexplained persistent fevers (in patients with either blood culture-positive or -negative endocarditis) despite optimal antibiotic therapy may reflect uncontrolled infection and warrant surgery. Surgical treatment is also advised for endocarditis caused by those organisms against which clinical experience indicates that effective antimicrobial therapy is lacking. This category includes infections caused by yeasts, fungi, *P. aeruginosa*, other highly resistant gram-negative bacilli, *Brucella* species, and probably *C. burnetii*.

S. AUREUS ENDOCARDITIS Mortality rates for *S. aureus* prosthetic valve endocarditis exceed 70% with medical treatment but are reduced to 25% with surgical treatment. In patients with intracardiac complications associated with *S. aureus* prosthetic valve

infection, surgical treatment reduces mortality by twentyfold. Surgical treatment should be considered for patients with *S. aureus* native aortic or mitral valve infection who haveTTE-demonstrable vegetations and remain septic during the initial week of therapy. Isolated tricuspid valve endocarditis, even with persistent fever, rarely requires surgery.

PREVENTION OF SYSTEMIC EMBOLI Mortality and persisting morbidity due to emboli are largely limited to patients suffering occlusion of cerebral or coronary arteries. Echocardiographic determination of vegetation size and anatomy, although predictive of patients at high risk of systemic emboli, does not identify those patients in whom the benefits of surgery to prevent emboli clearly exceed the risks of the surgical procedure. Net benefits favoring surgery are most likely when the risk of embolism is high and other surgical benefits can be achieved simultaneously -- e.g., repair of a moderately dysfunctional valve or debridement of a paravalvular abscess.

Timing of Cardiac Surgery Surgery to correct valvular dysfunction and progressive congestive heart failure should not be delayed simply to permit additional antibiotic therapy, since this course of action increases the risk of mortality. Similarly, surgery should not be delayed when the indication is uncontrolled or perivalvular infection. Delay is justified only when infection is controlled and congestive heart failure is fully compensated with medical therapy. Recrudescent endocarditis involving a prosthetic valve follows surgery in 2% of patients with culture-positive native valve endocarditis and 15% of patients with active prosthetic valve endocarditis. These risks are more acceptable than the high mortality rates that result when surgery is inappropriately delayed or not performed.

Among patients who have experienced a neurologic complication of endocarditis, further neurologic deterioration can occur as a consequence of cardiac surgery. The risk of significant neurologic exacerbation is related to the interval between the complication and surgery. Where feasible, cardiac surgery should be delayed for 2 to 3 weeks after a nonhemorrhagic embolic stroke and for 4 weeks after a hemorrhagic embolic stroke. A ruptured mycotic aneurysm should be clipped and cerebral edema allowed to resolve prior to cardiac surgery.

Extracardiac Complications Splenic abscess develops in 3 to 5% of patients with endocarditis. Effective therapy requires either computed tomography-guided percutaneous drainage or splenectomy. Mycotic aneurysms occur in 2 to 15% of endocarditis patients; half of these cases involve the cerebral arteries and present as headaches, focal neurologic symptoms, or hemorrhage. Cerebral aneurysms should be monitored by angiography. Some will resolve, but those that persist, enlarge, or leak should be treated surgically if possible. Extracerebral aneurysms present as local pain, a mass, local ischemia, or bleeding; generally these aneurysms are treated by resection.

OUTCOME

The outcome of infective endocarditis is affected by a variety of factors, some of which are interrelated. Factors with an adverse impact include older age, severe comorbid conditions, delayed diagnosis, involvement of prosthetic valves or the aortic valve, an invasive (*S. aureus*) or antibiotic-resistant (*P. aeruginosa*, yeast) pathogen, intracardiac

complications, and major neurologic complications. Death and poor outcome often are related not to failure of antibiotic therapy but rather to the interactions of comorbidities and endocarditis-related end-organ complications. The overall survival rate for patients with native valve endocarditis caused by viridans streptococci, HACEK organisms, or enterococci (susceptible to synergistic therapy) ranges from 85 to 90%. For *S. aureus* native valve endocarditis in patients who do not inject drugs, survival rates are 55 to 70%, whereas 85 to 90% of injection drug users survive this infection. Prosthetic valve endocarditis beginning within 2 months of valve replacement results in mortality rates of 40 to 50%, whereas rates are only 10 to 20% in later-onset cases.

PREVENTION

Antibiotics have been administered in conjunction with selected procedures considered to entail a risk for bacteremia and endocarditis. The benefits of antibiotic prophylaxis are not established and in fact may be modest: only 50% of patients with native valve endocarditis knew that they had a valve lesion predisposing to infection, most endocarditis cases do not follow a procedure, and 35% of cases are caused by organisms not targeted by prophylaxis. Dental treatments, the procedures most widely accepted as predisposing to endocarditis, are no more frequent during the 3 months preceding this diagnosis than in uninfected matched controls. Nevertheless, an expert committee of the American Heart Association, along with similar advisory groups in other developed countries, has identified procedures that may precipitate bacteremia with organisms that cause endocarditis (Table 126-6), patients who should receive prophylaxis based on the relative risk for developing endocarditis and the severity of subsequent infection (Table 126-7), and regimens that may be used for prophylaxis (Table 126-8). Except for an isolated secundum atrial septal defect and a totally corrected patent ductus arteriosus, ventricular septal defect, or pulmonary stenosis, patients with congenital heart defects continue to experience high rates of endocarditis despite total surgical correction of the defect. In vulnerable patients, maintaining good dental hygiene and aggressively treating local infections may reduce the risk of endocarditis.

(Bibliography omitted in Palm version)

Back to Table of Contents

127. INFECTIOUS COMPLICATIONS OF BITES AND BURNS - *Lawrence C. Madoff*

The skin is an essential component of the nonspecific immune system, protecting the host from potential pathogens in the environment. Breaches in this protective barrier thus represent a form of immunocompromise that predisposes the patient to infection. Bites and scratches from animals and humans allow the inoculation of microorganisms past the skin's protective barrier into deeper, susceptible host tissues. Thermal burns may cause massive destruction of the integument as well as derangements in humoral and cellular immunity, enabling environmental opportunists and components of the host's own skin flora to cause infection.

ANIMAL BITES AND SCRATCHES

Each year in the United States, between 1 and 2 million animal-bite wounds are sustained; the vast majority are inflicted by pet dogs and cats, which number more than 100 million. Other bite wounds are a consequence of encounters with animals in the wild or in occupational settings. While many of these wounds require minimal or no therapy, a significant number result in infection, which may be life-threatening. The microbiology of bite-wound infections in general reflects the oropharyngeal flora of the biting animal, although organisms from the soil, the skin of the animal and victim, and the animal's feces may also be involved.

Dog Bites Dogs are responsible for approximately 80% of bite wounds, an estimated 15 to 20% of which become infected. A study for the period 1992 through 1994 found that dog bites resulted in more than 900 emergency department visits each day in the United States. Most dog bites are provoked and are inflicted by the victim's pet or by a dog known to the victim. These bites frequently occur during efforts to break up a dogfight. Victims tend to be male, and bites most often involve a lower extremity. Infection typically manifests 8 to 24 h after the bite as pain at the site of injury with cellulitis accompanied by purulent, sometimes foul-smelling discharge. Septic arthritis and osteomyelitis may develop if the canine tooth penetrates synovium or bone. Systemic manifestations such as fever, lymphadenopathy, and lymphangitis may also occur. The microbiology of dog-bite wound infections is usually mixed and includesa-hemolytic streptococci, *Pasteurella* spp., *Staphylococcus* spp., *Eikenella corrodens*, and *Capnocytophaga canimorsus* (formerly designated DF-2). Many wounds also include anaerobic bacteria such as *Actinomyces*, *Fusobacterium*, *Prevotella*, and *Porphyromonas* spp.

While most infections resulting from dog-bite injuries are localized to the area of injury, many of the microorganisms involved are capable of causing systemic infection, including bacteremia, meningitis, brain abscess, endocarditis, and chorioamnionitis. These infections are particularly likely in hosts with edema or compromised lymphatic drainage in the involved extremity (e.g., following a bite on the arm after radical or modified radical mastectomy) and in patients who are immunocompromised by medication or disease (e.g., glucocorticoid use, systemic lupus erythematosus, acute leukemia, or hepatic cirrhosis). In addition, dog bites and scratches may result in systemic illnesses such as rabies (Chap. 197) and tetanus (Chap. 143).

Infection with *C. canimorsus* following dog-bite wounds may result in fulminant sepsis,

disseminated intravascular coagulation, and renal failure, particularly in hosts who have impaired hepatic function, who have undergone splenectomy, or who are immunosuppressed. This organism is a thin gram-negative rod that is difficult to culture on most solid media but grows in a variety of liquid media. The bacteria are occasionally seen within polymorphonuclear leukocytes on Wright-stained smears of peripheral blood from septic patients.

Cat Bites Although less common than dog bites, cat bites and scratches result in infection in more than half of all cases. Because the narrow, sharp feline incisors penetrate deeply into tissue, cat bites are more likely than dog bites to cause septic arthritis and osteomyelitis; the development of these conditions is particularly likely when punctures are located over or near a joint, especially in the hand. Women sustain cat bites more frequently than do men. These bites most often involve the hands and arms. Both bites and scratches from cats are prone to infection from organisms in the cat's oropharynx. *Pasteurella multocida*, a normal component of the feline oral flora, is a small gram-negative coccobacillus implicated in the majority of cat-bite wound infections. Like that of dog-bite wound infections, however, the microflora of cat-bite wound infections is usually mixed. Other microorganisms causing infection after cat bites are similar to those causing dog-bite wound infections.

The same risk factors for systemic infection following dog-bite wounds apply to cat-bite wounds. *Pasteurella* infections tend to advance rapidly, often within hours, causing severe inflammation accompanied by purulent drainage; *Pasteurella* may also be spread by respiratory droplets from animals, resulting in pneumonia or bacteremia. Like dog-bite wounds, cat-bite wounds may result in the transmission of rabies or in the development of tetanus. Infection with *Bartonella henselae* causes cat-scratch disease (Chap. 163) and is an important late consequence of cat bites and scratches. Tularemia (Chap. 161) has also been reported to follow cat bites.

Other Animal Bites Infections have been attributed to bites from many animal species, often as a consequence of occupational exposure (farmers, laboratory workers, veterinarians) or recreational exposure (hunters and trappers, wilderness campers, owners of exotic pets). Generally, the microflora of bite wounds reflects the oral flora of the biting animal. Most members of the cat family, including feral cats, harbor *P. multocida*. Bite wounds from aquatic animals such as alligators or piranhas may contain *Aeromonas hydrophila*. Venomous snakebites (Chap. 397) result in severe inflammatory responses and tissue necrosis -- conditions that render these injuries prone to infection. The snake's oral flora includes many species of aerobes and anaerobes, such as *Pseudomonas aeruginosa*, *Proteus* spp., *Staphylococcus epidermidis*, *Bacteroides fragilis*, and *Clostridium* spp. Bites from nonhuman primates are highly susceptible to infection with pathogens similar to those isolated from human bites (which are discussed later in this chapter). Bites from Old World monkeys (*Macaca*) may also result in the transmission of B virus (*Herpesvirus simiae*, cercopithecine herpesvirus), a cause of serious infection of the human central nervous system. Bites of seals, walruses, and polar bears may cause a chronic suppurative infection known as *seal finger*, which is probably due to one or more species of *Mycoplasma* colonizing these animals.

Small rodents, including rats, mice, and gerbils, as well as animals that prey on rodents may transmit *Streptobacillus moniliformis* (a microaerophilic, pleomorphic gram-negative

rod) or *Spirillum minor* (a spirochete), which cause a clinical illness known as *rat-bite fever*. The vast majority of cases in the United States are streptobacillary, whereas *Spirillum* infection occurs mainly in Asia.

In the United States, the risk of rodent bite mainly affects laboratory workers or inhabitants of rodent-infested dwellings (particularly children). Rat-bite fever is distinguished from acute bite-wound infection by its typical manifestation after the initial wound has healed. Streptobacillary disease follows an incubation period of 3 to 10 days. Fever, chills, myalgias, headache, and severe migratory arthralgias are usually followed by a maculopapular rash, which characteristically involves the palms and soles and may become confluent or purpuric. Complications include endocarditis, myocarditis, meningitis, pneumonia, and abscesses in many organs. *Haverhill fever* is an *S. moniliformis* infection acquired from contaminated milk or drinking water and has similar manifestations. Streptobacillary rat-bite fever was frequently fatal in the preantibiotic era. The differential diagnosis includes Rocky Mountain spotted fever, Lyme disease, leptospirosis, and secondary syphilis. The diagnosis is made by direct observation of the causative organisms in tissue or blood, by culture on enriched media, or by serologic testing with specific agglutinins.

Spirillum infection (referred to in Japan as *sodoku*) causes pain and purple swelling at the site of the initial bite, with associated lymphangitis and regional lymphadenopathy, after an incubation period of 1 to 4 weeks. The systemic illness includes fever, chills, and headache. The original lesion may eventually progress to an eschar. The infection is diagnosed by direct visualization of the spirochetes in blood or tissue or by animal inoculation.

Human Bites Human bites may be self-inflicted; may be sustained by medical personnel caring for patients; or may take place during fights, domestic abuse, or sexual activity. Human bites more frequently become infected than do bites inflicted by other animals. These infections reflect the diverse oral microflora of humans, which includes multiple species of aerobic and anaerobic bacteria. Common aerobic isolates include viridans streptococci, *Staphylococcus aureus*, *E. corrodens* (which is particularly common in clenched-fist injury; see below), and *Haemophilus influenzae*. Anaerobic species, including *Fusobacterium nucleatum* and *Prevotella*, *Porphyromonas*, and *Peptostreptococcus* spp., are isolated from 50% of human-bite wound infections; many of these isolates produceb-lactamases. The oral flora of hospitalized and debilitated patients often includes Enterobacteriaceae in addition to the usual organisms. Both HIV and hepatitis B virus have been reported to be transmitted by human bite, but these instances appear to be quite rare.

Human bites are categorized as "occlusional" injuries, which are inflicted by actual biting, and "clenched-fist" injuries, which are sustained when the fist of one individual strikes the teeth of another, causing traumatic laceration of the hand. For several reasons, clenched-fist injuries result in particularly serious infections. The deep spaces of the hand, including the bone, joint, and tendons, are frequently inoculated with organisms in the course of such injuries. The clenched position of the fist during injury, followed by extension of the hand, may further promote the introduction of bacteria as contaminated tendons retract beneath the skin's surface. Moreover, medical attention is often sought only after frank infection develops.

TREATMENT

Initial Assessment A careful history should be elicited, including the type of biting animal, the type of attack (provoked or unprovoked), and the amount of time elapsed since injury. Local and regional authorities should be contacted to determine whether an individual species could be rabid and/or to locate and observe the biting animal when rabies prophylaxis may be indicated (Chap. 197). Suspicious human-bite wounds should provoke careful questioning regarding domestic or child abuse. Details on antibiotic allergies, immunosuppression, splenectomy, liver disease, mastectomy, and immunization history should be obtained. The wound should be inspected carefully for evidence of infection, including redness, exudate, and foul odor. The type of wound (puncture, laceration, or scratch); the depth of penetration; and the possible involvement of joints, tendons, nerves, and bone should be assessed. It is often useful to include a diagram or photograph of the wound in the medical record. In addition, a general physical examination should be conducted and should include an assessment of vital signs as well as an evaluation for evidence of lymphangitis, lymphadenopathy, dermatologic lesions, and functional limitations. Injuries to the hand warrant consultation with a hand surgeon for the assessment of tendon, nerve, and muscular damage. Radiographs should be obtained when the bone may have been penetrated or a tooth fragment may be present. Culture and Gram's staining of all infected wounds are essential; anaerobic cultures should be undertaken if abscesses, devitalized tissue, or foul-smelling exudate is present. A small-tipped swab may be used to culture deep punctures or small lacerations. It is also reasonable to culture samples from uninfected wounds due to bites inflicted by animals other than dogs and cats, since the microorganisms causing disease are less predictable in these cases. A white blood cell count should be determined and blood cultured if systemic infection is suspected.

Wound Management Wound closure is controversial in bite injuries. Many authorities prefer not to attempt primary closure of wounds that are or may become infected, preferring to irrigate these wounds copiously, debride devitalized tissue, remove foreign bodies, and approximate the wound edges. Delayed primary closure may be undertaken after the risk of infection is over. Small uninfected wounds may be allowed to close by secondary intention. Puncture wounds due to cat bites should be left unsutured because of the high rate at which they become infected. Facial wounds are usually sutured after thorough cleaning and irrigation because of the importance of a good cosmetic result in this area and because anatomic factors such as an excellent blood supply and the absence of dependent edema lessen the risk of infection.

Antibiotic Therapy

Established Infection Antibiotics should be administered in all established bite-wound infections and should be chosen in light of the most likely potential pathogens, as indicated by the biting species and by Gram's stain and culture results (Table 127-1). For dog and cat bites, antibiotics should be effective against *S. aureus*, *Pasteurella* spp., *C. canimorsus*, streptococci, and oral anaerobes. For human bites, agents with activity against *S. aureus*, *H. influenzae*, andb-lactamase-positive oral anaerobes should be used. The combination of an extended-spectrum penicillin with ab-lactamase inhibitor (amoxicillin/clavulanic acid, ticarcillin/clavulanic acid, ampicillin/sulbactam) appears to

offer the most reliable coverage for these pathogens. Second-generation cephalosporins (cefuroxime, cefoxitin) also offer substantial coverage. The choice of antibiotics in penicillin-allergic patients (particularly those in whom immediate-type hypersensitivity makes the use of cephalosporins hazardous) is more difficult and is based primarily on in vitro sensitivity since data on clinical efficacy are inadequate. The combination of an antibiotic active against gram-positive cocci and anaerobes (such as clindamycin) with trimethoprim-sulfamethoxazole or a fluoroquinolone, which is active against many of the other potential pathogens, would appear reasonable. In vitro data suggest that either trovafloxacin or azithromycin alone provides coverage against most commonly isolated bite-wound pathogens.

Antibiotics are normally given for 10 to 14 days, but the response to therapy must be carefully monitored. Failure to respond should prompt a consideration of diagnostic alternatives and surgical evaluation for possible drainage or debridement. Complications such as osteomyelitis or septic arthritis mandate a longer duration of therapy.

Management of *C. canimorsus* sepsis requires a 2-week course of intravenous penicillin G (2 million units intravenously every 4 h) and supportive measures. Alternative agents for the treatment of *C. canimorsus* infection include cephalosporins and fluoroquinolones. Serious infection with *P. multocida* (e.g., pneumonia, sepsis, or meningitis) should also be treated with intravenous penicillin G. Alternative agents include second- or third-generation cephalosporins or ciprofloxacin.

Bites by venomous snakes may not require antibiotic treatment, but it is often difficult to distinguish signs of infection from tissue damage caused by the envenomation. Thus many authorities continue to recommend treatment directed against the snake's oral flora -- i.e., the administration of broadly active agents such as ceftriaxone (1 to 2 g intravenously every 12 to 24 h) or ampicillin/sulbactam (1.5 to 3.0 g intravenously every 6 h).

Seal finger appears to respond to doxycycline (100 mg twice daily for an interval guided by the response to therapy).

Presumptive or Prophylactic Therapy The use of antibiotics in patients presenting early after bite injury (within 8 h) is controversial. Although symptomatic infection will not yet be manifest in many of these wounds at this point, many early wounds will harbor pathogens, and many will become infected. Studies of the use of prophylactic antibiotics in wound infections are limited and have often included small numbers of cases in which various types of wounds have been managed according to various protocols. A recent meta-analysis of eight randomized trials of prophylactic antibiotics in patients with dog-bite wounds demonstrated a reduction of the rate of infection by approximately 50% with prophylaxis. However, in the absence of sound clinical trials, many clinicians base the decision to treat bite wounds with empirical antibiotics on the species of the biting animal; the location, severity, and extent of the bite wound; and the existence of comorbid conditions in the host. All human- and monkey-bite wounds should be treated presumptively because of the high rate of infection. Most cat-bite wounds, particularly those involving the hand, should be treated. Other factors favoring treatment for bite wounds include severe injury, as in crush wounds; potential bone or joint involvement; involvement of the hands or genital region; host immunocompromise, including that due

to liver disease or splenectomy; and prior mastectomy on the side of an involved upper extremity. When prophylactic antibiotics are administered, they are usually given for 3 to 5 days.

Rabies and Tetanus Prophylaxis Rabies prophylaxis, consisting of both passive administration of rabies immune globulin (with as much of the dose as possible infiltrated in and around the wound) and active immunization with rabies vaccine, should be given in consultation with local and regional public health authorities for many wild-animal (and some domestic-animal) bites and scratches as well as for certain nonbite exposures (Chap. 197). Rabies is endemic in a variety of animals, including dogs and cats in many areas of the world. Many local health authorities require the reporting of all animal bites. A tetanus booster immunization should be given if the patient has undergone primary immunization but has not received a booster dose in the past 5 years. Patients who have not previously completed primary immunization should be immunized and should also receive tetanus immune globulin. Elevation of the site of injury is an important adjunct to antimicrobial therapy. Immobilization of the infected area, especially the hand, is also beneficial.

BURNS

Epidemiology More than 2 million burn injuries are brought to medical attention in the United States each year. While many burn injuries are minor and require little or no intervention, approximately 70,000 persons are hospitalized for these injuries, and 20,000 of this number are burned severely enough to require admission to a specialized burn unit. Scalds, structural fires, and flammable liquids and gases are the major causes of burns, but electrical, chemical, and smoking-related sources are also important. Burns predispose to infection by damaging the protective barrier function of the skin, thus facilitating the entry of pathogenic microorganisms, and by inducing systemic immunosuppression. It is therefore not surprising that infectious complications are the major cause of morbidity and mortality in serious burn injury and that as many as 10,000 patients in the United States die of burn-related infections each year.

Pathophysiology Loss of the cutaneous barrier facilitates entry of the patient's own flora and of organisms from the hospital environment into the burn wound. The wound often contains devitalized or frankly necrotic tissue that quickly becomes contaminated with bacteria. Invasive infection -- localized and/or systemic -- occurs when bacteria penetrate viable tissue, usually below the eschar. Streptococci and staphylococci were the predominant causes of burn-wound infection in the preantibiotic era and remain important pathogens at present. With the advent of antimicrobial agents, *P. aeruginosa* became a major problem in burn-wound management. As antibiotics more effective against *Pseudomonas* have become available, fungi (particularly *Candida albicans*, *Aspergillus* spp., and the agents of mucormycosis) have emerged as increasingly important pathogens in burn-wound patients. Herpes simplex virus infection has also been found in burn wounds, especially on the face.

The frequency of infection parallels the extent and severity of burn injury. Severe burns cause defects in both cellular and humoral immunity that have a major impact on infection. For example, decreases in the number and activity of circulating helper T cells, increases in suppressor T cells, and diminution in levels of immunoglobulin follow

major burns. Neutrophil function has also been shown to be impaired after burns. The increased levels of multiple cytokines detected in burn patients are compatible with the widely held belief that the inflammatory response becomes dysregulated in these individuals. Increased permeability of the gut wall to bacteria and their components, such as endotoxin, also contributes to immune dysregulation and sepsis. Thus, the burn patient is predisposed to infection at remote sites (see below) as well as at the sites of burn injury.

Clinical Manifestations Since clinical indications of wound infection are difficult to interpret, wounds must be monitored carefully for changes that may reflect infection. A margin of erythema frequently surrounds the sites of burns and by itself is not usually indicative of infection. Signs of infection include the conversion of a partial-thickness to a full-thickness burn, color changes (e.g., the appearance of a dark brown or black discoloration of the wound), the new appearance of erythema or violaceous edema in normal tissue at the wound margins, the sudden separation of the eschar from subcutaneous tissues, and the degeneration of the wound with the appearance of a new eschar. The appearance of a green discoloration of the wound or subcutaneous fat or the development of ecthyma gangrenosum at a remote site points to a diagnosis of invasive *P. aeruginosa* infection. Changes in body temperature, hypotension, tachycardia, altered mentation, neutropenia or neutrophilia, thrombocytopenia, and renal failure may result from invasive burn wounds and sepsis. However, because profound alterations in homeostasis occur as a consequence of burns per se and because inflammation without infection is a normal component of these injuries, the assessment of these changes is complicated. Alterations in body temperature, for example, are attributable to thermoregulatory dysfunction; tachycardia and hyperventilation accompany the metabolic changes induced by extensive burn injury and are not necessarily indicative of bacterial sepsis.

Given the difficulty of evaluating burn wounds solely on the basis of clinical observation and laboratory data, wound biopsies are necessary for definitive diagnosis of infection. The timing of these biopsies can be guided by clinical changes, but in some centers burn wounds are routinely biopsied at regular intervals. The biopsy specimen is examined for histologic evidence of bacterial invasion, and quantitative microbiologic cultures are performed. The presence of $>10^5$ viable bacteria per gram of tissue is highly suggestive of invasive infection and of a dramatically increased risk of sepsis. Histopathologic evidence of invasion of viable tissue by microorganisms is a more definitive indicator of infection. A blood culture positive for the same organism seen in large quantities in biopsied tissue is a reliable indicator of burn sepsis. Surface cultures may provide some indication of the microorganisms present in the hospital environment but are not indicative of the etiology of infection.

In addition to infection of the burn wound itself, a number of other infections due to the immunosuppression caused by extensive burns and the manipulations necessary for clinical care put burn patients at risk. Pneumonia, now the most common infectious complication among hospitalized burn patients, is most often nosocomially acquired via the respiratory route; septic pulmonary emboli may also occur. Suppurative thrombophlebitis may complicate the vascular catheterization necessary for fluid and nutritional support in burns. Endocarditis, urinary tract infection, bacterial chondritis (particularly in patients with burned ears), and intraabdominal infection also complicate

serious burn injury.

TREATMENT

The ultimate goal of burn-wound management is closure and healing of the wound. Early surgical excision of burned tissue, with extensive debridement of necrotic tissue and grafting of skin or skin substitutes, greatly decreases the mortality associated with severe burns. In addition, the three widely used topical antimicrobial agents -- silver sulfadiazine cream, mafenide acetate cream, and silver nitrate -- dramatically decrease the bacterial burden of burn wounds and reduce the incidence of burn-wound infection; they are routinely applied to partial- and full-thickness burns. All three agents are broadly active against many bacteria and against some fungi and are useful before bacterial colonization is established. Silver sulfadiazine is often used initially, but its value can be limited by bacterial resistance. Mafenide acetate has broader activity; the cream penetrates eschars and thus can prevent or treat infection beneath the eschars. The foremost disadvantages of this agent are that it can inhibit carbonic anhydrase, resulting in metabolic acidosis, and that it elicits hypersensitivity reactions in up to 7% of patients. This agent is most often used when gram-negative bacteria invade the burn wound and when treatment with silver sulfadiazine fails.

When invasive wound infection is diagnosed, topical therapy should be changed to mafenide acetate. Subeschar clysis (the direct instillation of an antibiotic, often piperacillin, under the eschar into wound tissues) is a useful adjunct to surgical and systemic antimicrobial therapy. Systemic treatment with antibiotics active against the pathogens present in the wound should be instituted. In the absence of culture data, treatment is broad and should cover organisms commonly encountered in the particular burn unit. Usually such coverage is achieved with an antibiotic active against gram-positive pathogens, such as oxacillin (2 g intravenously every 4 h), and with antibiotics active against *P. aeruginosa* and other gram-negative rods, such as mezlocillin (3 g intravenously every 4 h) and gentamicin (5 mg/kg intravenously per day). In the penicillin-allergic patient, vancomycin (1 g intravenously every 12 h) may be substituted for oxacillin (and is efficacious when methicillin-resistant *S. aureus* is present), and ciprofloxacin (400 mg intravenously every 12 h) may be substituted for mezlocillin. Patients with burn wounds frequently have alterations in metabolism and renal clearance mechanisms that mandate the monitoring of serum antibiotic levels; the levels achieved with standard doses are often subtherapeutic.

In general, prophylactic systemic antibiotics have no role in the management of burn wounds (except for minor burns in outpatients) and can in fact lead to colonization with resistant microorganisms. An exception involves cases requiring burn-wound manipulation. Since procedures such as debridement, excision, or grafting frequently result in bacteremia, prophylactic systemic antibiotics are administered at the time of burn-wound manipulation; the particular agents used should be chosen on the basis of data obtained by wound culture or data on the hospital's resident flora. All burn-injury patients should undergo tetanus booster immunization if they have completed primary immunization but have not received a booster dose in the past 5 years. Patients without prior immunization should receive tetanus immune globulin and undergo primary immunization. Infection control measures play a major role in preventing burn-wound infection and limiting the spread of antibiotic-resistant nosocomial pathogens.

(Bibliography omitted in Palm version)

Back to Table of Contents

128. INFECTIONS OF THE SKIN, MUSCLE, AND SOFT TISSUES - *Dennis L. Stevens*

ANATOMIC RELATIONSHIPS: CLUES TO THE DIAGNOSIS OF SOFT TISSUE INFECTIONS

Protection against infection of the epidermis is dependent on the mechanical barrier afforded by the stratum corneum, since the epidermis itself is devoid of blood vessels (Fig. 128-1). Disruption of this layer by burns or bites (Chap. 127), abrasions, foreign bodies, primary dermatologic disorders (e.g., herpes simplex, varicella, and ecthyma gangrenosum), surgery, or vascular or pressure ulcer allows penetration of bacteria to the deeper structures. Similarly, the hair follicle can serve as a portal either for components of the normal flora (e.g., *Staphylococcus*) or for extrinsic bacteria (e.g., *Pseudomonas* in hot-tub folliculitis). Intracellular infection of the squamous epithelium with vesicle formation may arise from cutaneous inoculation, as in infection with herpes simplex virus (HSV) type 1; from the dermal capillary plexus, as in varicella and infections due to other viruses associated with viremia; or from cutaneous nerve roots, as in herpes zoster. Bacteria infecting the epidermis, such as *Streptococcus pyogenes*, may be translocated laterally to deeper structures via lymphatics, an event that results in the rapid superficial spread of erysipelas. Later, engorgement or obstruction of lymphatics causes flaccid edema of the epidermis, another characteristic of erysipelas.

The rich plexus of capillaries beneath the dermal papillae provides nutrition to the stratum germinativum, and physiologic responses of this plexus produce important clinical signs and symptoms. For example, infective vasculitis of the plexus results in petechiae, Osler's nodes (Fig. 126-CD3), Janeway lesions (Fig. 19-CD2), and palpable purpura, which are important clues to the existence of endocarditis (Chap. 126). In addition, metastatic infection within this plexus can result in cutaneous manifestations of disseminated fungal infection (Chap. 205), gonococcal infection (Chap. 147), *Salmonella* infection (Chap. 156), *Pseudomonas* infection (i.e., ecthyma gangrenosum; Chap. 155), meningococcemia (Chap. 146), and staphylococcal infection (Chap. 139). The plexus also provides access for bacteria to the circulation, thereby facilitating local spread or bacteremia. The postcapillary venules of this plexus are a major site of polymorphonuclear leukocyte sequestration, diapedesis, and chemotaxis to the site of cutaneous infection.

Exaggeration of these physiologic mechanisms by excessive levels of cytokines or bacterial toxins causes leukostasis, venous occlusion, and pitting edema. Edema with purple bullae, ecchymosis, and cutaneous anesthesia suggests loss of vascular integrity and necessitates exploration of the deeper structures for evidence of necrotizing fasciitis or myonecrosis. An early diagnosis requires a high level of suspicion in instances of unexplained fever and of pain and tenderness in the soft tissue, even in the absence of acute cutaneous inflammation.

INFECTIONS ASSOCIATED WITH VESICLES (Table 128-1)

Vesicle formation due to infection is caused by viral proliferation within the epidermis. In varicella and variola, viremia precedes the onset of a diffuse centripetal rash that progresses from macules to vesicles, then to pustules, and finally to scabs over the

course of 1 to 2 weeks. Vesicles of varicella have a "dewdrop" appearance and develop in crops randomly about the trunk, extremities, and face over 3 to 4 days. Herpes zoster occurs in a single dermatome; the appearance of vesicles is preceded by pain for several days. Zoster may occur in persons of any age but is most common among immunosuppressed individuals and elderly patients, whereas most cases of varicella occur in young children. Vesicles due toHSV are found on or around the lips (HSV-1) or genitals (HSV-2) but may appear on the head and neck of young wrestlers (herpes gladiatorum) or on the digits of health care workers (herpetic whitlow). Coxsackievirus A16 characteristically causes vesicles on the hands, feet, and mouth of children. Orf is caused by a DNA virus related to smallpox virus and infects the fingers of individuals who work around goats and sheep. Molluscum contagiosum virus induces flaccid vesicles on the skin of healthy and immunocompromised individuals.

Rickettsialpox begins following mite-bite inoculation of *Rickettsia akari* into the skin. A papule with a central vesicle evolves to form a 1- to 2.5-cm painless crusted black eschar with an erythematous halo and proximal adenopathy. While more common in the northeastern United States and the Ukraine in 1940-1950, rickettsialpox has recently been described in Ohio, Arizona, and Utah. Blistering dactylitis is a painful, vesicular, localized *Staphylococcus aureus* or group A streptococcal infection of the pulps of the distal digits of the hands.

INFECTIONS ASSOCIATED WITH BULLAE (Table 128-1)

Staphylococcal scalded-skin syndrome (SSSS) (Fig. 18-CD5) in neonates is caused by a toxin (exfoliatin) from phage group II *S. aureus*. SSSS must be distinguished from toxic epidermal necrolysis (TEN), which occurs primarily in adults, is drug-induced, and has a higher mortality. Punch biopsy with frozen section is useful in making this distinction since the cleavage plane is the stratum corneum in SSSS (Fig. 128-1) and the stratum germinativum in TEN. Intravenous g-globulin is a promising treatment for TEN. Necrotizing fasciitis and gas gangrene also induce bulla formation (see "Necrotizing Fasciitis," below). Halophilic vibrio infection can be as aggressive and fulminant as necrotizing fasciitis; a helpful clue in its diagnosis is a history of exposure to waters of the Gulf of Mexico or the Atlantic seaboard or (in a patient with cirrhosis) the ingestion of raw seafood. The etiologic organism (*Vibrio vulnificus*) is highly susceptible to tetracycline.

INFECTIONS ASSOCIATED WITH CRUSTED LESIONS (Table 128-1)

Impetigo contagiosa is caused by *S. pyogenes*, and bullous impetigo is due to *S. aureus* (Fig. 128-CD1). Both skin lesions may have an early bullous stage but then appear as thick crusts with a golden-brown color (Fig. 128-CD2). Streptococcal lesions are most common among children 2 to 5 years of age, and epidemics may occur in settings of poor hygiene, particularly among children of lower socioeconomic status in tropical climates. It is important to recognize impetigo contagiosa because of its relationship to poststreptococcal glomerulonephritis. Superficial dermatophyte infection (ringworm) can occur on any skin surface, and skin scrapings with KOH staining are diagnostic. Primary infections with dimorphic fungi such as *Blastomyces dermatitidis* and *Sporothrix schenckii* can initially present as crusted skin lesions resembling ringworm. Disseminated infection with *Coccidioides immitis* can also involve the skin, and biopsy

and culture should be performed on crusted lesions in patients from endemic areas. Crusted nodular lesions caused by *Mycobacterium chelonei* have been described in HIV-seropositive patients. Treatment with clarithromycin looks promising.

FOLLICULITIS (Table 128-1)

Hair follicles serve as a portal for a number of bacteria, though *S. aureus* is the most common cause of localized folliculitis (Fig. 128-CD3). Sebaceous glands empty into hair follicles and ducts and, if blocked, form sebaceous cysts, which may resemble staphylococcal abscesses or may become secondarily infected. Infection of sweat glands (hidradenitis suppurativa) can also mimic infection of hair follicles, particularly in the axillae. Chronic folliculitis is uncommon except in acne vulgaris, where constituents of the normal flora (e.g., *Propionibacterium acnes*) may play a role.

Diffuse folliculitis occurs in two settings. "Hot-tub folliculitis" (Fig. 128-CD4) is caused by *Pseudomonas aeruginosa* in waters that are insufficiently chlorinated and maintained at temperatures between 37 and 40°C. Infection is usually self-limited, though bacteremia and shock have been reported. Swimmer's itch occurs when a skin surface is exposed to water infested with freshwater avian schistosomes. Warm water temperatures and alkaline pH are suitable for mollusks that serve as intermediate hosts between birds and humans. Free-swimming schistosomal cercariae readily penetrate human hair follicles or pores but quickly die and elicit a brisk allergic reaction causing intense itching and erythema.

PAPULAR AND NODULAR LESIONS (Table 128-1)

Raised lesions of the skin occur in many different forms. *Mycobacterium marinum* infections of the skin may present as cellulitis or as raised erythematous nodules. Erythematous papules are early manifestations of cat-scratch disease (primary site of inoculation) and bacillary angiomatosis (*Bartonella henselae*). Raised serpiginous or linear eruptions are characteristic of cutaneous larva migrans, which is caused by burrowing larvae of dog or cat hookworms (*Ancylostoma braziliense*) and which humans acquire through contact with soil that has been contaminated with dog or cat feces. Similar burrowing raised lesions are present in dracunculiasis caused by migration of the adult female nematode *Dracunculus medinensis*. Nodules caused by *Onchocerca volvulus* may range from 1 to 10 cm in diameter and occur largely in persons bitten by *Simulium* flies in Africa. The nodules contain the adult worm encased in fibrous tissue. Migration of microfilariae into the eyes may result in blindness. Verruca peruana is caused by *Bartonella bacilliformis*, which is transmitted to humans by the sandfly *Phlebotomus*. This condition can take the form of single gigantic lesions (several centimeters in diameter) or multiple small lesions (several millimeters in diameter). Numerous subcutaneous nodules may also be present in cysticercosis caused by larvae of *Taenia solium*. Multiple erythematous papules develop in schistosomiasis; each represents a cercarial invasion site. Skin nodules as well as thickened subcutaneous tissue are prominent features of lepromatous leprosy. Large nodules or gummas are features of tertiary syphilis (Fig. 128-CD5), whereas flat papulosquamous lesions are characteristic of secondary syphilis. Human papillomavirus may cause singular warts (verruca vulgaris; Fig. 128-CD6) or multiple warts in the anogenital area (condylomata acuminata).

ULCERS WITH OR WITHOUT ESCHARS (Table 128-1)

Cutaneous anthrax begins as a pruritic papule, which develops within days into an ulcer with surrounding vesicles and edema and then into an enlarging ulcer with a black eschar. Cutaneous anthrax may cause chronic nonhealing ulcers with an overlying dirty-gray membrane, though lesions may also mimic psoriasis, eczema, or impetigo. Ulceroglandular tularemia may have associated ulcerated skin lesions with painful regional adenopathy. Although buboes are the major cutaneous manifestation of plague, in 25% of cases ulcers with eschars, papules, or pustules are also present.

Mycobacterium ulcerans typically causes chronic skin ulcers on the extremities of individuals living in the tropics. *Mycobacterium leprae* may be associated with cutaneous ulcerations in patients with lepromatous leprosy related to Lucio's phenomenon or during reversal reactions. *Mycobacterium tuberculosis* may also cause ulcerations, papules, or erythematous macular lesions of the skin in both normal and immunocompromised patients.

Decubitus ulcers are due to tissue hypoxia secondary to pressure-induced vascular insufficiency and may become secondarily infected with components of the skin and gastrointestinal flora, including anaerobes. Ulcerative lesions on the anterior shins may be due to pyoderma gangrenosum, which must be distinguished from similar lesions of infectious etiology by histologic evaluation of biopsy sites. Ulcerated lesions on the genitals may be either painful (chancroid) or painless (primary syphilis).

ERYSIPELAS (Table 128-1)

Erysipelas is due to *S. pyogenes* and is characterized by an abrupt onset of fiery-red swelling of the face or extremities. The distinctive features of erysipelas are well-defined indurated margins, particularly along the nasolabial fold; rapid progression; and intense pain. Flaccid bullae may develop during the second or third day of illness, but extension to deeper soft tissues is rare. Treatment with penicillin is effective; swelling may progress despite appropriate treatment, though fever, pain, and the intense red color diminish. Desquamation of the involved skin occurs 5 to 10 days into the illness. Infants and elderly adults are most commonly afflicted, and the severity of systemic toxicity varies.

CELLULITIS (Table 128-1)

Cellulitis is an acute inflammatory condition of the skin that is characterized by localized pain, erythema, swelling, and heat. Cellulitis may be caused by indigenous flora colonizing the skin and appendages (e.g., *S. aureus* and *S. pyogenes*) or by a wide variety of exogenous bacteria. Because the exogenous bacteria involved in cellulitis occupy unique niches in nature, a thorough history including epidemiologic data provides important clues to etiology. When there is drainage, an open wound, or an obvious portal of entry, Gram's stain and culture provide a definitive diagnosis. In the absence of these findings, the bacterial etiology of cellulitis is difficult to establish, and in some cases staphylococcal and streptococcal cellulitis may have similar features. Even with needle aspiration of the leading edge or a punch biopsy of the cellulitis tissue itself,

cultures are positive in only 20% of cases. This observation suggests that relatively low numbers of bacteria may cause cellulitis and that the expanding area of erythema within the skin may be a direct effect of extracellular toxins or of the soluble mediators of inflammation elicited by the host.

Bacteria may gain access to the epidermis through cracks in the skin, abrasions, cuts, burns, insect bites, surgical incisions, and intravenous catheters. Cellulitis caused by *S. aureus* spreads from a central localized infection, such as an abscess, folliculitis, or an infected foreign body (e.g., a splinter, a prosthetic device, or an intravenous catheter). In contrast, cellulitis due to *Staphylococcus pyogenes* is a more rapidly spreading, diffuse process frequently associated with lymphangitis and fever. Recurrent streptococcal cellulitis of the lower extremities may be caused by organisms of group A, C, or G in association with chronic venous stasis or with saphenous venectomy for coronary artery bypass surgery. Streptococci also cause recurrent cellulitis among patients with chronic lymphedema resulting from elephantiasis, lymph node dissection, or Milroy's disease. Recurrent staphylococcal cutaneous infections are more common among individuals who have eosinophilia and elevated serum levels of IgE (Job's syndrome) and among nasal carriers of staphylococci. Cellulitis caused by *S. agalactiae* (group B streptococci) occurs primarily in elderly patients and those with diabetes mellitus or peripheral vascular disease. *Haemophilus influenzae* typically causes periorbital cellulitis in children in association with sinusitis, otitis media, or epiglottitis. It is unclear whether this form of cellulitis will (like meningitis) become less common as a result of the impressive efficacy of the *H. influenzae* type b vaccine.

Many other bacteria also cause cellulitis. Fortunately, these organisms occur in such characteristic settings that a good history provides useful clues to the diagnosis. Cellulitis associated with cat bites and, to a lesser degree, with dog bites is commonly caused by *Pasteurella multocida*, though in the latter case *Staphylococcus intermedius* and *Capnocytophaga canimorsus* (formerly DF-2) must also be considered. Sites of cellulitis and abscesses associated with dog bites and human bites also contain a variety of anaerobic organisms, including *Fusobacterium*, *Bacteroides*, aerobic and anaerobic streptococci, and *Eikenella corrodens*. *Pasteurella* is notoriously resistant to dicloxacillin and nafcillin but is sensitive to all other b-lactam antimicrobials as well as to quinolones, tetracycline, and erythromycin. Ampicillin/clavulanate, ampicillin/sulbactam, and cefoxitin are good choices for the treatment of animal or human bite infections. *Aeromonas hydrophila* causes aggressive cellulitis in tissues surrounding lacerations sustained in fresh water (lakes, rivers, and streams). This organism remains sensitive to aminoglycosides, fluoroquinolones, chloramphenicol, trimethoprim-sulfamethoxazole, and third-generation cephalosporins; it is resistant to ampicillin, however.

P. aeruginosa causes three types of soft tissue infection: ecthyma gangrenosum in neutropenic patients, hot-tub folliculitis, and cellulitis following penetrating injury. Most commonly, *P. aeruginosa* is introduced into the deep tissues when a person steps on a nail. Treatment includes surgical inspection and drainage, particularly if the injury also involves bone or joint capsule. Choices for empirical treatment while antimicrobial susceptibility data are awaited include an aminoglycoside, a third-generation cephalosporin (ceftazidime, cefoperazone, or cefotaxime), a semisynthetic penicillin (ticarcillin, mezlocillin, or piperacillin), or a fluoroquinolone (though drugs of the last class are not indicated for the treatment of children <13 years old).

Gram-negative bacillary cellulitis, including that due to *P. aeruginosa*, is most common among hospitalized, immunocompromised hosts. Cultures and sensitivity tests are critically important in this setting because of multidrug resistance (Chap. 155).

The gram-positive aerobic rod *Erysipelothrix rhusiopathiae* is most often associated with fish and domestic swine and causes cellulitis primarily in bone renderers and fishmongers (Fig. 128-CD7). *E. rhusiopathiae* remains susceptible to most b-lactam antibiotics (including penicillin), erythromycin, clindamycin, tetracycline, and cephalosporins but is resistant to sulfonamides, chloramphenicol, and vancomycin. Its resistance to vancomycin, which is unusual among gram-positive bacteria, is of potential clinical significance since this agent is sometimes used in empirical therapy for skin infection. Fish food containing the water flea *Daphnia* is sometimes contaminated with *M. marinum*, which can cause cellulitis or granulomas on skin surfaces exposed to the water in aquariums or injured in swimming pools. Rifampin plus ethambutol has been an effective therapeutic combination in some cases, though no comprehensive studies have been undertaken. In addition, some strains of *M. marinum* are susceptible to tetracycline or to trimethoprim-sulfamethoxazole.

NECROTIZING FASCIITIS (Table 128-1,Fig. 18-CD4)

Necrotizing fasciitis, formerly called streptococcal gangrene, may be associated with group A *Streptococcus* or mixed aerobic-anaerobic bacteria or may occur as part of gas gangrene caused by *Clostridium perfringens*. Early diagnosis may be difficult when pain or unexplained fever is the only presenting manifestation. Swelling then develops and is followed by brawny edema and tenderness. With progression, dark red induration of the epidermis appears along with bullae filled with blue or purple fluid. Later the skin becomes friable and takes on a bluish, maroon, or black color. By this stage, thrombosis of blood vessels in the dermal papillae (Fig. 128-1) is extensive. Extension of infection to the level of the deep fascia causes this tissue to take on a brownish-gray appearance. Rapid spread occurs along fascial planes, through venous channels and lymphatics. Patients in the later stages are toxic and frequently manifest shock and multiorgan failure.

Necrotizing fasciitis caused by mixed aerobic-anaerobic bacteria begins with a breach in the integrity of a mucous membrane barrier, such as the mucosa of the gastrointestinal or genitourinary tract. The portal can be a malignancy, diverticulum, hemorrhoid, anal fissure, or urethral tear. Other predisposing factors include peripheral vascular disease, diabetes mellitus, surgery, and penetrating injury to the abdomen. Leakage into the perineal area results in a syndrome called *Fournier's gangrene*, characterized by massive swelling of the scrotum and penis with extension into the perineum or the abdominal wall and legs.

Necrotizing fasciitis caused by *S. pyogenes* has increased in frequency and severity since 1985. It frequently begins deep at the site of a nonpenetrating minor trauma such as a bruise or a muscle strain. Seeding of the site via transient bacteremia is likely, though most patients deny antecedent streptococcal infection. Alternatively, *S. pyogenes* may reach the deep fascia from a site of cutaneous infection or penetrating trauma. Toxicity is severe, and renal impairment may precede the development of

shock. In 20 to 40% of cases, myositis occurs concomitantly, and, as in gas gangrene (see below), serum creatinine phosphokinase values may be markedly elevated. Necrotizing fasciitis due to mixed aerobic-anaerobic bacteria may be associated with gas in the deep tissue, but gas is not usually present when the cause is *S. pyogenes*. Prompt surgical exploration down to the deep fascia and muscle is essential. Necrotic tissue must be surgically removed, and Gram's staining and culture of excised tissue are useful in establishing whether group A streptococci, mixed aerobic-anaerobic bacteria, or *Clostridium* spp. are present (see "Treatment" below).

MYOSITIS/MYONECROSIS (Table 128-1)

Muscle involvement can occur with virus infection (e.g., influenza virus, dengue virus, or coxsackievirus B) or parasitic invasion (e.g., trichinosis, cysticercosis, or toxoplasmosis). Although myalgia can occur in most of these infections, severe muscle pain is the hallmark of pleurodynia (coxsackie virus B), trichinosis, and bacterial infection. Acute rhabdomyolysis predictably occurs with clostridial and streptococcal myositis but may also be associated with influenza virus, echovirus, coxsackievirus, Epstein-Barr virus, and *Legionella* infection.

Pyomyositis is usually due to *S. aureus*, is common in tropical areas, and generally has no known portal of entry. Infection remains localized, and shock does not develop unless organisms produce toxic shock syndrome toxin 1 or certain enterotoxins and the patient lacks antibodies to the toxin produced by the infecting organisms. In contrast, *S. pyogenes* may induce primary myositis referred to as *streptococcal necrotizing myositis*, which is associated with severe systemic toxicity. Myonecrosis occurs concomitantly with necrotizing fasciitis in about 50% of cases. Both are part of the streptococcal toxic shock syndrome.

Gas gangrene usually follows severe penetrating injuries that result in interruption of the blood supply and introduction of soil into wounds. Such cases of traumatic gangrene are usually caused by *C. perfringens*, *C. septicum*, or *C. histolyticum*. Rarely, latent or recurrent gangrene can occur years after penetrating trauma, most likely owing to dormant spores that reside at the site of previous injury. Spontaneous nontraumatic gangrene among patients with neutropenia, gastrointestinal malignancy, diverticulosis, or recent radiation therapy to the abdomen is caused by several clostridial species, although *C. septicum* is most common. The tolerance of this anaerobe to oxygen probably explains why it can initiate infection spontaneously in normal tissue anywhere in the body.

Synergistic nonclostridial anaerobic myonecrosis, also known as necrotizing cutaneous myositis and synergistic necrotizing cellulitis, is a variant of necrotizing fasciitis caused by mixed aerobic and anaerobic bacteria with the exclusion of clostridial organisms (see "Necrotizing Fasciitis," above).

DIAGNOSIS

This chapter has emphasized the physical appearance and location of lesions within the soft tissues as important diagnostic clues. The temporal progression of the lesions as well as the patient's travel history, animal exposure or bite history, age, underlying

disease status, and lifestyle are also crucial considerations in the formulation of a narrowed differential diagnosis. However, even the astute clinician may find it challenging to diagnose all infections of the soft tissues by history and inspection alone. Soft tissue radiography, computed tomography, and magnetic resonance imaging may be useful in determining the depth of infection and should be performed in patients with rapidly progressing lesions or in those with evidence of systemic inflammatory response syndrome. These tests are particularly valuable for defining a localized abscess or detecting gas in tissue. Unfortunately, they may reveal only soft tissue swelling and thus are not specific for fulminant infections such as necrotizing fasciitis or myonecrosis caused by group A *Streptococcus*, where gas is not found in lesions.

Aspiration of the leading edge or punch biopsy with frozen section may be helpful if the results are positive, but false-negative results occur in approximately 80% of cases. There is some evidence that aspiration alone may be superior to injection and aspiration using normal saline. Frozen sections are especially useful in distinguishingSSSSfrom TEN and are quite valuable in cases of necrotizing fasciitis. Open surgical inspection with debridement as indicated is clearly the best way to determine the extent and severity of infection and to obtain material for Gram's staining and culture. Such an aggressive approach is important and may be lifesaving if undertaken early in the course of fulminant infections where there is evidence of systemic toxicity.

TREATMENT

A full description of the treatment of all the clinical entities described herein is beyond the scope of this chapter. As a guide to the clinician in selecting appropriate treatment, the antimicrobial agents useful in the most common and the most fulminant cutaneous infections are listed inTable 128-2.

Early and aggressive surgical exploration is essential in patients with suspected necrotizing fasciitis, myositis, or gangrene in order to (1) visualize the deep structures, (2) remove necrotic tissue, (3) reduce compartment pressure, and (4) obtain suitable material for Gram's staining and for aerobic and anaerobic cultures. Appropriate empirical antibiotic treatment for mixed aerobic-anaerobic infections could consist of ampicillin/sulbactam, cefoxitin, or the following combination: (1) clindamycin (600 to 900 mg intravenously every 8 h) or metronidazole (750 mg every 6 h) plus (2) ampicillin or ampicillin/sulbactam (2 to 3 g intravenously every 6 h) plus (3) gentamicin (1.0 to 1.5 mg/kg every 8 h). Group A streptococcal and clostridial infection of the fascia and/or muscle carries a mortality rate of 20 to 50% with penicillin treatment. In experimental models of streptococcal and clostridial necrotizing fasciitis/myositis, clindamycin has exhibited markedly superior efficacy, but no comparative trials have been performed in humans. Hyperbaric oxygen treatment may also be useful in gas gangrene due to clostridial species. Antibiotic treatment should be continued until all signs of systemic toxicity have resolved, all devitalized tissue has been removed, and granulation tissue has developed (Chaps. 140,145, and167).

In summary, infections of the skin and soft tissues are diverse in presentation and severity and offer a great challenge to the clinician. This chapter provides an approach to diagnosis and understanding of the pathophysiologic mechanisms involved in these infections. More in-depth information is found in chapters on specific infections.

(Bibliography omitted in Palm version)

Back to Table of Contents

129. OSTEOMYELITIS - *James H. Maguire*

Osteomyelitis, an infection of bone, is caused most commonly by pyogenic bacteria and mycobacteria. Classification of cases on the basis of the causative agent; the route, duration, and anatomic location of infection; and local and systemic host factors provides a useful framework for evaluating the patient and planning treatment.

PATHOGENESIS AND PATHOLOGY

Microorganisms enter bone by the hematogenous route, by direct introduction from a contiguous focus of infection, or by a penetrating wound. Trauma, ischemia, and foreign bodies enhance the susceptibility of bone to microbial invasion by exposing sites to which bacteria can bind. Phagocytes attempt to contain the infections and, in the process, release enzymes that lyse bone. Pus spreads into vascular channels, raising intraosseous pressure and impairing the flow of blood; as the untreated infection becomes chronic, ischemic necrosis of bone results in the separation of large devascularized fragments (*sequestra*). When pus breaks through the cortex, subperiosteal or soft tissue abscesses form, and the elevated periosteum deposits new bone (the *involucrum*) around the sequestrum. Bacteria escape host defenses by adhering tightly to damaged bone, by entering and persisting within osteoblasts, and by coating themselves and underlying surfaces with a protective polysaccharide-rich biofilm.

Microorganisms, infiltrates of neutrophils, and congested or thrombosed blood vessels are the principal histologic findings of acute osteomyelitis. The distinguishing feature of chronic osteomyelitis is necrotic bone, which is characterized by the absence of living osteocytes. Mononuclear cells predominate in chronic infections, and granulation and fibrous tissues replace bone that has been resorbed by osteoclasts. In the chronic stage, organisms may be too few to be seen.

HEMATOGENOUS OSTEOMYELITIS

Hematogenous infection accounts for ~20% of cases of osteomyelitis and primarily affects children, in whom the long bones are infected, and older adults and intravenous drug users, in whom the spine is the usual site of infection.

Acute Hematogenous Osteomyelitis Infection usually involves a single bone, most commonly the tibia, femur, or humerus. Bacteria settle in the well-perfused metaphysis, where functioning phagocytes are scarce, a network of venous sinusoids slows the flow of blood, and fenestrations in capillaries allow organisms to escape into the extravascular space. Because vascular anatomy changes with age, hematogenous infection of long bones is uncommon during adulthood and, when it occurs, usually involves the diaphysis.

In children, the source of bacteremia is often inapparent, although there may have been recent blunt trauma to the extremity leading to a small intraosseous hematoma or vascular obstruction. On presentation, the child usually appears acutely ill, with high fever, chills, localized pain and tenderness, and leukocytosis. Cutaneous erythema and swelling indicate extension of pus through the cortex. During infancy and after puberty,

infection may spread through the epiphysis into the joint space. In children of other ages (i.e., between infancy and puberty), extension of infection through the cortex results in involvement of joints if the metaphysis is intracapsular. Thus, septic arthritis of the elbow, shoulder, and hip may complicate osteomyelitis of the proximal radius, humerus, and femur, respectively.

Plain radiographs initially show soft tissue swelling, but the first change in bone -- a periosteal reaction -- is not evident until at least 10 days after the onset of infection. Lytic changes can be detected after 2 to 6 weeks, when 50 to 75% of bone density has been lost. Rarely, a well-circumscribed lytic lesion, or *Brodie's abscess*, is seen in a child who has been in pain for several months but has had no fever.

Chronic Hematogenous Osteomyelitis With prompt treatment,<5% of cases of acute hematogenous osteomyelitis progress to chronic osteomyelitis. On average, 10 days are required for the formation of necrotic bone, but plain radiographs are unable to detect sequestra or sclerotic new bone for many weeks.

A protracted clinical course, long periods of quiescence, and recurrent exacerbations are characteristic of chronic osteomyelitis. Sinus tracts between bone and skin may drain purulent material and occasionally pieces of necrotic bone. An increase in drainage, pain, or the erythrocyte sedimentation rate (ESR) signals an exacerbation. Fever is unusual except when obstruction of a sinus tract leads to soft tissue infection. Rare late complications include pathologic fractures, squamous cell carcinoma of the sinus tract, and amyloidosis.

Vertebral Osteomyelitis Organisms reach the well-perfused vertebral body of adults via spinal arteries and quickly spread from the end plate into the disk space and then to the adjacent vertebral body. The infection may originate in the urinary tract, and it does so particularly often among elderly men. Other sources of bacteremia include endocarditis, soft tissue infection, and a contaminated intravenous line; these sources are usually obvious. Diabetes mellitus, hemodialysis, and intravenous drug use carry an increased risk of spinal infection. Penetrating injuries and surgical procedures to the spine may cause nonhematogenous vertebral osteomyelitis or infection localized to the disk.

Most patients with vertebral osteomyelitis report neck or back pain; 15% describe atypical pain in the chest, the abdomen, or an extremity that is due to irritation of nerve roots. Symptoms are localized to the lumbar spine more often than to the thoracic spine (>50% vs. 35% of cases) or the cervical spine in pyogenic infections, but the thoracic spine is involved most commonly in tuberculous spondylitis (Pott's disease). Percussion over the involved vertebra elicits tenderness, and physical examination may reveal spasm of the paraspinal muscles and a limitation of motion. More than 50% of patients experience a subacute illness in which a vague, dull pain gradually intensifies over 2 to 3 months; fever is low grade or absent, and the white blood cell count is normal. An acute presentation with high fever and toxicity is less common and suggests ongoing bacteremia.

Usually, by the time the patient seeks medical attention, theESR is elevated, and plain radiographs show irregular erosions in the end plates of adjacent vertebral bodies and

narrowing of the intervening disk space. This radiographic pattern is virtually diagnostic of bacterial infection because tumors and other diseases of the spine rarely cross the disk space. Computed tomography (CT) or magnetic resonance imaging (MRI) may demonstrate epidural, paraspinal, retropharyngeal, mediastinal, retroperitoneal, or psoas abscesses that originate in the spine. An epidural abscess may evolve suddenly or over several weeks; irreversible paralysis may result from failure to recognize the classic clinical presentation of a spinal epidural abscess: spinal pain progressing to radicular pain and weakness.

Microbiology More than 95% of cases of hematogenous osteomyelitis are caused by a single organism. *Staphylococcus aureus* accounts for 50% of isolates. Other common pathogens include group B streptococci and *Escherichia coli* during the newborn period and group A streptococci and *Haemophilus influenzae* in early childhood. Vertebral osteomyelitis is due to *E. coli* and other enteric bacilli in ~25% of cases. *S. aureus*, *Pseudomonas aeruginosa*, and *Serratia* infections are associated with intravenous drug use in some parts of the United States and may involve the sacroiliac, sternoclavicular, or pubic joints as well as the spine. *Salmonella* spp. and *S. aureus* are the major causes of long-bone osteomyelitis complicating sickle cell anemia and other hemoglobinopathies. Tuberculosis and brucellosis affect the spine more often than other bones. Other common sites of tuberculous osteomyelitis include the small bones of the hands and feet, the metaphyses of long bones, the ribs, and the sternum.

Unusual causes of hematogenous osteomyelitis include disseminated histoplasmosis, coccidioidomycosis, and blastomycosis in endemic areas. Immunocompromised persons on rare occasions develop osteomyelitis due to atypical mycobacteria, *Bartonella henselae*, or *Pneumocystis carinii* or to species of *Candida*, *Cryptococcus*, or *Aspergillus*. Syphilis, yaws, varicella, and vaccinia may involve bone. The etiology of chronic relapsing multifocal osteomyelitis, an inflammatory condition of children that is characterized by recurrent episodes of painful lytic lesions in multiple bones, has not yet been identified.

OSTEOMYELITIS SECONDARY TO A CONTIGUOUS FOCUS OF INFECTION

Clinical Features This broad category of osteomyelitis includes infections introduced by penetrating injuries and surgical procedures and by direct extension of infection from adjacent soft tissues. It accounts for the greatest number of cases of osteomyelitis and occurs most commonly in adults.

Frequently, the diagnosis is not made until the infection has already become chronic. The pain, fever, and inflammatory signs due to acute osteomyelitis may be attributed to the original injury or soft tissue infection. An indolent infection may become apparent only weeks or months later, when a sinus tract develops, a surgical wound breaks down, or a fracture fails to heal. It may be impossible to distinguish radiographic abnormalities due to osteomyelitis from those due to the precipitating condition.

A special type of contiguous-focus osteomyelitis occurs in the setting of peripheral vascular disease and nearly always involves the small bones of the feet of adult diabetic patients. Diabetic neuropathy exposes the foot to frequent trauma and pressure sores, and the patient may be unaware of infection as it spreads into bone. Poor tissue

perfusion impairs normal inflammatory responses and wound healing and creates a milieu that is conducive to anaerobic infections. It is often during the evaluation of a nonhealing ulcer, a swollen toe, or acute cellulitis that a radiograph provides the first evidence of osteomyelitis. If bone is palpable during examination of the base of an ulcer with a blunt surgical probe, osteomyelitis is likely.

Microbiology *S. aureus* is a pathogen in more than half of cases of contiguous-focus osteomyelitis. However, in contrast to hematogenous osteomyelitis, these infections are often polymicrobial and are more likely to involve gram-negative and anaerobic bacteria. Hence a mixture of staphylococci, streptococci, enteric organisms, and anaerobic bacteria may be isolated from a diabetic foot infection or pelvic osteomyelitis underlying a decubitus ulcer. Aerobic and anaerobic bacteria cause osteomyelitis following surgery or soft tissue infection of the oropharynx, paranasal sinuses, gastrointestinal tract, or female genital tract. *S. aureus* is the principal cause of postoperative infections; coagulase-negative staphylococci are common pathogens after implantation of orthopedic appliances; and these organisms as well as gram-negative enteric bacilli, atypical mycobacteria, and *Mycoplasma* may cause sternal osteomyelitis after cardiac surgery. Infection with *P. aeruginosa* is frequently associated with puncture wounds of the foot or with thermal burns, and *Pasteurella multocida* infection commonly follows cat bites (Chap. 127).

DIAGNOSIS

Early diagnosis of acute osteomyelitis is critical because prompt antibiotic therapy may prevent the necrosis of bone. The evaluation usually begins with plain radiographs because of their ready availability, although they frequently show no abnormalities during early infection. TheESR and C-reactive protein levels are elevated in most cases of active osteomyelitis, including those in which constitutional symptoms and leukocytosis are lacking. These findings are not specific to osteomyelitis, however, and the ESR is occasionally normal in early infections. In 95% of cases, the technetium radionuclide scan using ^{99m}Tc diphosphonate is positive within 24 h of the onset of symptoms. Falsely negative scans usually indicate obstruction of blood flow to the bone. Because the uptake of technetium reflects osteoblastic activity and skeletal vascularity, the bone scan cannot differentiate osteomyelitis from fractures, tumors, infarction, or neuropathic osteopathy.^{67}Ga citrate- and^{111}In-labeled leukocyte or immunoglobulin scans, which have greater specificity for inflammation, may help distinguish infectious from noninfectious processes and indicate inflammatory changes within bones that for other reasons are already abnormal on radiography and technetium scanning. Ultrasound can be used to diagnose osteomyelitis by the detection of subperiosteal fluid collections, soft tissue abscesses adjacent to bone, and periosteal thickening and elevation.

MRIis as sensitive as the bone scan for the diagnosis of acute osteomyelitis because it is able to detect changes in the water content of marrow. MRI yields better anatomic resolution of epidural abscesses and other soft tissue processes than CTand is currently the imaging technique of choice for vertebral osteomyelitis (Fig. 129-1).

The role of diagnostic imaging in chronic osteomyelitis is to determine the presence of active infection and delineate the extent of debridement necessary to remove necrotic

bone and abnormal soft tissues. Although plain films accurately reflect chronic changes, theCT scan is more sensitive for the detection of sequestra, sinus tracts, and soft tissue abscesses. Both CT and ultrasound are useful for guiding percutaneous aspiration of subperiosteal and soft tissue fluid collections. Sequential technetium and gallium or indium scans may help determine whether infection is active and may distinguish infection from noninflammatory bone changes; these methods do not, however, provide good anatomic detail.MRIprovides detailed information about the activity and the anatomic extent of infection but does not always distinguish osteomyelitis from healing fractures and tumors. MRI is particularly useful in distinguishing cellulitis from osteomyelitis in the diabetic foot; however, no imaging modality consistently distinguishes infection from neuropathic osteopathy.

Appropriate samples for microbiologic studies should be obtained in all cases of suspected osteomyelitis before the initiation of antimicrobial therapy. Blood cultures are indicated in acute cases and are positive in more than one-third of cases of hematogenous osteomyelitis in children and in 25% of cases of vertebral osteomyelitis in adults. If the clinical picture demands immediate antibiotic therapy or if blood cultures are negative, samples from needle aspiration of pus in bone or soft tissues or from a bone biopsy should be obtained for culture.

The results of culture of specimens obtained by swabbing of a sinus tract or the base of an ulcer correlate poorly with the organisms infecting the bone. For this reason, in cases of chronic osteomyelitis and contiguous-focus osteomyelitis, samples for aerobic and anaerobic culture should be obtained from several sites by percutaneous needle aspiration, percutaneous biopsy, or intraoperative biopsy at the time of debridement. Isolates of coagulase-negative staphylococci and other organisms of low virulence should not automatically be disregarded as contaminants, especially in the presence of prosthetic materials. Special culture media may be necessary for the isolation of mycobacteria, fungi, and less common pathogens. In some cases, histopathologic examination of biopsy specimens may be the only way to make a diagnosis.

TREATMENT

Antibiotic Therapy Antibiotics are administered only after appropriate specimens have been obtained for culture. The antibiotics selected should be bactericidal and, at least initially, should be given intravenously. When necessary, empirical therapy is guided by findings on Gram's staining of a specimen from the bone or abscess or is chosen to cover the most likely pathogens. Empirical therapy in most cases should include high doses of an agent active against *S. aureus* (such as oxacillin, nafcillin, a cephalosporin, or vancomycin) and, if gram-negative organisms are likely to be involved, a third-generation cephalosporin, an aminoglycoside, or a fluoroquinolone.

Specific intravenous therapy is based on the in vitro susceptibility of the organism(s) isolated from bone or blood. Penicillin G (3 to 4 million units every 4 h) is the drug of choice for the treatment of infections due to penicillin-sensitive staphylococci and streptococci; nafcillin or oxacillin (2 g every 4 h) is preferred for penicillin-resistant, methicillin-sensitive staphylococci. Cefazolin (1 to 2 g every 8 h) or vancomycin [15 mg/kg (up to 1 g) every 12 h] is an alternative for persons allergic to penicillins. Infections due to methicillin-resistant staphylococci are treated with vancomycin.

Regimens for infections due to susceptible gram-negative rods include ampicillin (2 g every 4 h), cefazolin, a second-generation cephalosporin such as cefuroxime (1.5 g every 8 h), or a fluoroquinolone such as ciprofloxacin (400 mg every 12 h) or levofloxacin (500 mg every 24 h). Initial therapy for osteomyelitis due to *P. aeruginosa* or *Enterobacter* spp. should not consist of ab-lactam antibiotic alone because of the potential for these organisms to develop resistance during therapy. Appropriate intravenous therapies for *P. aeruginosa* infections include tobramycin (1.7 mg/kg every 8 h, or 5 to 7 mg/kg every 24 h) and a broad-spectrumb-lactam compound such as ticarcillin (3 g every 4 h), ceftazidime (1 to 2 g every 8 h), or aztreonam (1 to 2 g every 8 h); a fluoroquinolone may be substituted for one of the latter. *Enterobacter* infections can be treated with a fluoroquinolone alone or with combinations of a broad-spectrum b-lactam antibiotic and gentamicin in the same doses as tobramycin. Serum levels of aminoglycosides should be monitored closely to avoid toxicity.

The duration of therapy is typically 4 to 6 weeks; at-home intravenous administration of antibiotics or oral therapy is appropriate for motivated and medically stable patients. Antibiotics that require infrequent dosing, such as ceftriaxone and vancomycin, facilitate home therapy. Children with acute hematogenous osteomyelitis routinely receive oral antibiotics after 5 to 10 days of parenteral therapy if signs of active infection have resolved; such treatment has been as successful as standard parenteral therapy. The doses of oral penicillins or cephalosporins required for the treatment of osteomyelitis are several times higher than the doses of these drugs given for common infections. Adults may not tolerate these high doses as well as children, and, except in the case of the fluoroquinolones, few data support the use of oral antibiotics by adults. For treatment of osteomyelitis due to Enterobacteriaceae, oral administration of an agent such as ciprofloxacin (750 mg every 12 h) or levofloxacin (500 mg every 24 h) has been as successful as intravenous administration of b-lactam antibiotics. Caution should be exercised in the use of fluoroquinolones as the sole agents for treatment of infection due to *S. aureus* or *P. aeruginosa* because resistance may develop during therapy. Addition of rifampin to a quinolone has yielded encouraging results in infections due to *S. aureus*, but further studies are necessary to confirm these findings. Oral administration of clindamycin (300 to 450 mg every 6 h) or metronidazole (500 mg every 8 h) results in high drug levels in serum and can take the place of intravenous regimens for the treatment of *Bacteroides* infections. Oral clindamycin has produced good results in therapy for osteomyelitis due to *S. aureus*, especially in children. There are few data to support the routine use of the serum minimal bactericidal concentration (MBC) other than to document adherence to treatment.

Acute Osteomyelitis Early treatment of acute hematogenous osteomyelitis of childhood with 4 to 6 weeks of an appropriate antibiotic is usually successful; treatment for<3 weeks has resulted in a 10-fold greater rate of failure. Surgical intervention in childhood cases is indicated for intraosseous or subperiosteal abscesses, concomitant septic arthritis, and failure of the acute signs of infection to improve in 24 to 48 h. Acute hematogenous osteomyelitis of bones other than the spine in adults often requires surgical debridement.

Vertebral Osteomyelitis A 4- to 6-week course of treatment with an appropriate antibiotic is usually sufficient to cure vertebral osteomyelitis. Failure of theESR to drop by two-thirds or more of its pretreatment level is an indication for longer treatment.

Surgery is seldom necessary, even in cases of many months' duration, except in instances of spinal instability, new or progressive neurologic deficits, large soft tissue abscesses that cannot be drained percutaneously, or a failure of medical treatment. Patients should maintain bed rest until back pain has declined to the point at which ambulation is possible. Body casts are no longer used. Spontaneous fusion of involved vertebrae occurs in the majority of cases after successful treatment.

Contiguous-Focus Osteomyelitis Even when diagnosed early, contiguous-focus osteomyelitis usually requires surgery in addition to 4 to 6 weeks of appropriate antibiotic therapy because of underlying soft tissue infection or damage to bone from an injury or surgery. A 2-week course of antibiotics following thorough debridement and soft tissue coverage has yielded excellent results in treatment of superficial osteomyelitis involving only the outer cortex of bone.

Chronic Osteomyelitis The risks and benefits of aggressive therapy for chronic osteomyelitis should be weighed before any attempt is made to eradicate the infection. Some patients with extensive disease prefer to live with their infections rather than undergo multiple surgical procedures, take prolonged courses of antimicrobial therapy, and face the risk of loss of an extremity. Such persons often benefit from intermittent courses of oral antibiotics to suppress acute exacerbations.

Once the decision has been made to treat chronic osteomyelitis aggressively, the patient's nutritional and metabolic status should be optimized to expedite healing of soft tissues and bone. Antibiotic administration should be started several days before surgery to reduce inflammation if the etiology of the infection is known preoperatively. If not, antibiotic therapy should be withheld until surgical debridement. An empirical antibiotic regimen is started intraoperatively after culture specimens are obtained. A 4- to 6-week course of appropriate antibiotic therapy is given postoperatively on the basis of the susceptibility pattern of organisms isolated from the bone. The benefit of prolonged oral antibiotic therapy after 4 to 6 weeks of parenteral therapy remains unproven. There is insufficient information to recommend the routine use of hyperbaric oxygen to enhance the killing of microorganisms by phagocytes or of instillation pumps and antibiotic-impregnated methacrylate beads to deliver high levels of antibiotics to the bone.

The success of therapy for chronic osteomyelitis rests largely on the complete surgical removal of necrotic bone and abnormal soft tissues. Modern imaging techniques allow accurate preoperative delineation of tissues to be debrided, but it remains difficult for the surgeon to determine intraoperatively whether all necrotic and infected tissue has been removed. In the past, the inability to repair large defects in bone and soft tissue limited the extent of debridement. Muscle flaps and skin grafts are now used routinely to cover large soft tissue defects and fill dead space, and bone grafts and vascularized bone transfer may restore a seriously compromised bone to a functional state.

In infections of recent fractures, internal fixators are often left in place, and the infection is controlled by limited debridement and suppressive antibiotic therapy. Definitive surgical/antimicrobial therapy is delayed until after bony union of the fracture is achieved. If there is nonunion of the fracture or loosening of the fixator, the appliance should be removed, the bone debrided, and an external fixator or a new internal fixator

applied.

Osteomyelitis of the small bones of the feet in persons with vascular disease also requires surgical treatment. The effectiveness of the surgery is limited by the blood supply to the site and the body's ability to heal the wound. Revascularization of the extremity is indicated if the vascular disease involves large arteries. In cases of decreased perfusion due to small-vessel disease, foot-sparing surgery may fail, and the best option is suppressive therapy or amputation. The duration of antibiotic therapy depends on the surgical procedure performed. When the infected bone is removed entirely but residual infection of soft tissues remains, antibiotic therapy should be given for 2 weeks; if amputation eliminates infected bone and soft tissue, standard surgical prophylaxis is given; otherwise, postoperative antibiotics must be given for 4 to 6 weeks.

(Bibliography omitted in Palm version)

Back to Table of Contents

130. INTRAABDOMINAL INFECTIONS AND ABSCESSES - *Dori F. Zaleznik*, *Dennis L. Kasper*

Intraperitoneal infections generally arise because a normal anatomic barrier is disrupted. This disruption may occur when the appendix, a diverticulum, or an ulcer ruptures; when the bowel wall is weakened by ischemia, tumor, or inflammation (e.g., in inflammatory bowel disease); or with adjacent inflammatory processes, such as pancreatitis or pelvic inflammatory disease, in which enzymes (in the former case) or organisms (in the latter) may leak into the peritoneal cavity. Whatever the inciting event, once inflammation develops and organisms usually contained within the bowel or another organ enter the normally sterile peritoneal space, a predictable series of events takes place. Intraabdominal infections occur in two stages: peritonitis and -- if it goes untreated -- abscess formation. The types of microorganisms predominating in each stage of infection are responsible for the pathogenesis of disease.

PERITONITIS

The peritoneal cavity is large but is divided into compartments. The upper and lower peritoneal cavities are divided by the transverse mesocolon; the greater omentum extends from the transverse mesocolon and from the lower pole of the stomach to line the lower peritoneal cavity. The pancreas, duodenum, and ascending and descending colon are located in the anterior retroperitoneal space; the kidneys, ureters, and adrenals are found in the posterior retroperitoneal space. The other organs, including liver, stomach, gallbladder, spleen, jejunum, ileum, transverse and sigmoid colon, cecum, and appendix, are found within the peritoneal cavity itself. Normally the cavity is lined with a serous membrane that can serve as a conduit for fluids -- a property utilized in peritoneal dialysis. A small amount of fluid, sufficient to allow movement of organs, is normally present in the peritoneal space. This fluid is serous, with a protein content (consisting mainly of albumin) of<30 g/L and fewer than 300 white blood cells (WBCs, generally mononuclear cells) per microliter. In the presence of infection, some of these compartments collect fluid or pus more often than others. These compartments include the pelvis (the lowest portion), the subphrenic spaces on the right and left sides, and Morrison's pouch, which is a posterosuperior extension of the subhepatic spaces and is the lowest part of the paravertebral groove when a patient is recumbent. The falciform ligament separating the right and left subphrenic spaces appears to act as a barrier to the spread of infection; consequently, it is unusual to find bilateral subphrenic collections.

SPONTANEOUS BACTERIAL PERITONITIS

Peritonitis is either primary (without an apparent source of contamination) or secondary. The types of organisms found and the clinical presentation of these two processes are different. In adults, primary or spontaneous bacterial peritonitis (SBP) occurs most commonly in conjunction with cirrhosis of the liver (frequently the result of alcoholism). It virtually always develops in patients with ascites. Nevertheless, it is not a common event, occurring in£10% of cirrhotic patients. The cause of SBP has not been established definitively but is believed to involve hematogenous spread of organisms in a patient in whom a diseased liver and altered portal circulation result in a defect in the usual filtration function. Organisms are able to multiply in ascites, a good medium for

growth. The proteins of the complement cascade have been found in peritoneal fluid, with lower levels in cirrhotic patients than in patients with ascites of other etiologies. The opsonic and phagocytic properties of neutrophils are decreased in patients with advanced liver disease.

The presentation ofSBPdiffers from that of secondary peritonitis. The most common manifestation is fever, which is reported in as many as 80% of patients. Ascites is found but virtually always predates infection. Abdominal pain, an acute onset of symptoms, and peritoneal irritation detected during physical examination can be helpful diagnostically, but the absence of any of these findings does not exclude this often-subtle diagnosis. It is vital to sample the peritoneal fluid of any cirrhotic patient with ascites and fever. The finding of>300 polymorphonuclear leukocytes (PMNs) per microliter is diagnostic for SBP, according to Conn. The microbiology of SBP is also distinctive. While enteric gram-negative bacilli such as *Escherichia coli* are most commonly encountered, gram-positive organisms such as streptococci, enterococci, or even pneumococci are sometimes found. In SBP, a single organism is typically isolated; anaerobes are found less frequently in SBP than in secondary peritonitis, in which a mixed flora including anaerobes is the rule. In fact, if SBP is suspected and multiple organisms including anaerobes are recovered from the peritoneal fluid, the diagnosis must be reconsidered and a source of secondary peritonitis sought.

The diagnosis ofSBP is not easy. It depends on the exclusion of a primary intraabdominal source of infection. Contrast-enhanced computed tomography (CT) is very useful in identifying an intraabdominal source for infection. It may be difficult to recover organisms from cultures of peritoneal fluid, presumably because the burden of organisms is low. However, the yield can be improved if 10 mL of peritoneal fluid is placed directly into a blood culture bottle. Bacteremia frequently accompanies SBP; therefore, blood should be cultured simultaneously. No specific radiographic studies are helpful in the diagnosis of SBP. A plain film of the abdomen would be expected to show ascites. Chest and abdominal radiography should be performed in patients with abdominal pain to exclude free air, which signals a perforation.

TREATMENT

Treatment for SBPis directed at the isolate from blood or peritoneal fluid. Gram's staining of peritoneal fluid often gives negative results in primary peritonitis; therefore, until culture results become available, empirical therapy should cover gram-negative aerobic bacilli and gram-positive cocci. Ampicillin plus gentamicin is a reasonable initial regimen. Third-generation cephalosporins, carbapenems, or broad-spectrum penicillin/b-lactamase inhibitor combinations are also options. Empirical coverage for anaerobes is not necessary. After the infecting organism is identified, therapy should be narrowed to target that specific pathogen. Patients with SBP usually respond within 72 h to appropriate antibiotic therapy.

SECONDARY PERITONITIS

Secondary peritonitis develops when bacteria contaminate the peritoneum as a result of spillage from an intraabdominal viscus. The organisms found almost always constitute a mixed flora in which facultative gram-negative bacilli and anaerobes predominate,

especially when the contaminating source is colonic. Early in the course of infection, when the host response is directed toward containment of the infection, exudate containing fibrin andPMNs is found. Early death in this setting is attributable to gram-negative bacillary sepsis and to potent endotoxins circulating in the bloodstream (Chap. 124). Gram-negative bacilli, particularly *E. coli*, are common bloodstream isolates, but *Bacteroides fragilis* bacteremia occurs as well. The severity of abdominal pain and the clinical course depend on the inciting process. The species of organisms isolated from the peritoneum also vary with the source of the initial process and the normal flora present at that site. Peritonitis can result primarily from chemical irritation or bacterial contamination. For example, as long as the patient is not achlorhydric, a ruptured gastric ulcer will release low-pH gastric contents that will serve as a chemical irritant. The normal flora of the stomach comprises the same organisms found in the oropharynx (Chap. 167) but in lower numbers. The surfaces of teeth contain ~10^7aerobic and 10^7anaerobic organisms per milliliter of saliva; the normally acidic stomach contains an equal ratio of aerobic and anaerobic species, but in concentrations more in the range of 10^5/mL. After meals, when gastric acidity is highest, this number may fall to 10^3/mL. Thus, the bacterial burden in a ruptured gastric ulcer -- or even a duodenal ulcer -- is negligible compared with that in a ruptured appendix. The normal flora of the colon below the ligament of Treitz contains ~10^{11}anaerobic organisms per gram of feces but only 10^8aerobes per gram; therefore, anaerobic species account for 99% of the bacteria. Leakage of colonic contents (pH 7 to 8) does not cause significant chemical peritonitis, but infection is intense because of the heavy bacterial load.

Depending on the inciting event, local symptoms may initially be found in secondary peritonitis -- for example, epigastric pain from a ruptured gastric ulcer. In appendicitis (Chap. 291), the initial presenting symptoms are often vague, with periumbilical discomfort and nausea followed in a number of hours by pain more localized to the right lower quadrant. Unusual locations of the appendix (including a retrocecal position) can complicate this presentation further. Once infection has spread to the peritoneal cavity, however, pain increases, particularly with infection involving the parietal peritoneum, which is innervated extensively. Patients usually lie motionless, often with knees drawn up to avoid stretching the nerve fibers of the peritoneal cavity. Coughing and sneezing, which increase pressure within the peritoneal cavity, are associated with sharp pain. There may or may not be pain localized to the infected or diseased organ from which secondary peritonitis has arisen. Patients with secondary peritonitis generally have abnormal findings on abdominal examination, with marked voluntary and involuntary guarding of the anterior abdominal musculature. Later findings include tenderness, especially rebound tenderness. In addition, there may be localized findings in the area of the inciting event. In general, patients are febrile, with marked leukocytosis and a left shift of theWBCs to earlier granulocyte forms.

While recovery of organisms from peritoneal fluid is easier in secondary than in primary peritonitis, a tap of the abdomen is rarely the procedure of choice in secondary peritonitis. An exception is in cases involving trauma, where the possibility of a hemoperitoneum may need to be excluded early.

TREATMENT

Treatment for secondary peritonitis includes early administration of antibiotics aimed