ICSI

Institute for Clinical Systems Improvement

Technology Assessment Report

The information contained in this *ICSI Technology Assessment Report* is intended primarily for health professionals and the following expert audiences:

- physicians, nurses, and other health care professional and provider organizations;
- health plans, health systems, health care organizations, hospitals and integrated health care delivery systems;
- medical specialty and professional societies;
- researchers;
- federal, state and local government health care policy makers and specialists; and
- employee benefit managers.

This *ICSI Technology Assessment Report* should not be construed as medical advice or medical opinion related to any specific facts or circumstances. If you are not one of the expert audiences listed above you are urged to consult a health care professional regarding your own situation and any specific medical questions you may have. In addition, you should seek assistance from a health care professional in interpreting this *ICSI Technology Assessment Report* and applying it in your individual case.

This *ICSI Technology Assessment Report* is designed to assist clinicians by providing a scientific assessment, through review and analysis of medical literature, of the safety and efficacy of medical technologies and is not intended either to replace a clinician's judgment or to suggest that a particular technology is or should be a standard of medical care in any particular case. Standards of medical care are determined on the basis of all the facts and circumstances involved in a particular case and are subject to change as scientific knowledge and technology advance and practice patterns evolve.

Technology Assessment

Densitometry as a Diagnostic Tool for the Identification and Treatment of Osteoporosis in Women

Work Group Membership
Richard Kopher, MD, Work Group Leader
HealthPartners Medical Group
Beth Averbeck, MD
HealthPartners Medical Group
Bart Clarke, MD
Mayo Clinic
John Schousboe, MD
HealthSystem Minnesota

Prepared under the direction of the
Technology Assessment Committee
James C. Smith, M.D., Chairperson
Nancy L. Greer, Ph.D., Staff

Abstract

Description of Treatment/Procedure

Osteoporosis is the most common metabolic bone disorder. It is characterized by abnormalities in the amount and architectural arrangement of bone tissue which lead to impaired skeletal strength and susceptibility to fractures. In white women, osteoporosis is defined as a bone mineral density (BMD) more than 2.5 standard deviations below the mean for young, normal women. BMD is commonly measured with dual-energy x-ray absorptiometry (DXA), quantitative computed tomography (QCT), and quantitative ultrasound (QUS). Both central (hip, spine) and peripheral (wrist, calcaneus) sites may be tested.

Potential Uses

The National Osteoporosis Foundation recommends testing for BMD in the following groups **provided that the decision to test is based on an individual's risk profile and that test results could influence the treatment decision**:

a. all postmenopausal women <65 years who have one or more additional risk factors (besides menopause),
b. all women ≥65 years regardless of additional risk factors,
c. postmenopausal women who present with fractures,
d. women who are considering therapy for osteoporosis, if testing would facilitate the decision, and
e. women who have been on hormone replacement therapy for prolonged periods.

Contraindications

BMD testing is contraindicated if the results will not influence patient or physician behavior.

Efficacy of Procedure

Prediction of fracture risk is of primary interest. A recent study showed that a single measure of BMD (single-photon absorptiometry was used in this study) has a predictive value for fragility fractures for as long as 25 years. Recent advances in DXA technology including peripheral DXA and fan beam DXA have facilitated measurement of peripheral sites and reduced scan times, respectively. However, neither of these new techniques offers as high a precision as has been found with conventional DXA systems. In prospective studies, QUS results were found to be comparable to BMD in predicting hip and non-spine fractures in elderly women, however, there is no agreement on how to interpret the data for the purpose of diagnosing osteoporosis and the measurement is less precise than DXA.

Committee Summary

The ICSI Technology Assessment Committee finds that:

1. Osteoporosis is a significant health issue reducing quality of life and resulting in significant treatment costs
2. The value of preventive interventions has been proven. There is a range of possible interventions some of which can be initiated without BMD screening. The available options should be thoroughly explained to patients.
3. SPA, SXA, DPA, DXA, QCT, and QUS are safe procedures.
4. DXA is most commonly used because it allows measurements of the spine and hip and offers higher precision than other systems. Higher precision allows for repeat testing to assess the effects of treatment. QUS is more portable and less expensive but less precise than DXA. QUS may be used to encourage women toward therapy. (Conclusion Grade II)
5. **There is insufficient evidence to support mass screening for BMD; the need for BMD testing must be determined on an individual patient basis**. Testing is of value when making individual decisions about therapies in lieu of estrogen replacement therapy as well as when an individual's decision about estrogen replacement therapy would be influenced by her knowledge of her BMD.
6. The value of repetitive testing for monitoring the effects of therapy or for making decisions about interventions in elderly women is poorly understood. Yearly densitometry is not clinically indicated. Decisions about when to repeat testing must be individualized based on the patient's health status, risk status, and baseline BMD.
7. Each clinical center should establish their own short-term precision values.

Approved January, 2000 **TA # 31 (Revised)**

Institute for Clinical Systems Improvement

Technology Assessment Report

Densitometry as a Diagnostic Tool for the Identification and Treatment of Osteoporosis in Women

Work Group Members

Work Group Leader:
Richard Kopher, MD
Obstetrics/Gynecology
HealthPartners Medical Group

Beth Averbeck, MD
Internal Medicine
HealthPartners Medical Group

Bart Clarke, MD
Endocrinology
Mayo Clinic

John Schousboe, MD
Endocrinology
HealthSystem Minnesota

Prepared under the direction of the
Technology Assessment Committee
James C. Smith, M.D., Chairman
Nancy L. Greer, Ph.D., Staff

TA#31 (Revised)
Approved January, 2000

Technology Assessment Committee

James Smith, M.D., Chair
HealthPartners Medical Group

Kenneth Engberg, M.D.
HealthPartners Medical Group

Gwen Halaas, M.D.
HealthPartners Medical Group

John Haugen, M.D.
HealthSystem Minnesota

George Isham, M.D.
HealthPartners

Brad Johnson, M.D.
HealthEast Clinics

George Klee, M.D.
Mayo Clinic

Richard Kopher, M.D.
HealthPartners Medical Group

James Lillehei, M.D.
Aspen Medical Group (Retired)

George Logan, M.D.
HealthSystem Minnesota

Kirsten Hall Long, Ph.D.
Mayo Clinic

James Mickman, M.D.
HealthPartners Medical Group

James Milavetz, M.D.
HealthSystem Minnesota

Lorre Ochs, M.D.
HealthPartners Medical Group

Mark Paller, M.D.
University of Minnesota

Howard Stang, M.D.
HealthPartners Medical Group

Seth Wolpert, M.D.
HealthPartners Medical Group

ICSI Technology Assessment Report Process

- A topic is selected by the Technology Assessment Committee based on expression of interest in that topic from the ICSI medical groups and HealthPartners.
- A work group of 4 to 6 physicians and other health care professionals who are experts in the topic area is assembled (with a formally designated leader).
- The literature search is completed using MEDLINE and PREMEDLINE; in addition, bibliographies of articles obtained from the literature search are examined to identify articles that may have been missed and work group members are asked to provide key references. The evidence is graded according to the system described in the reference section of the report.
- The ICSI staff person prepares a draft report.
- The work group meets to review the draft report and directs the ICSI staff person in revising the report
- The work group leader presents the report to the ICSI Technology Assessment Committee. Committee members review the report to determine whether the conclusions are supported by the evidence cited. The Committee often requests revisions prior to approving the report for review and comment.
- The report is distributed to the ICSI Medical Groups for review and comment. Comments received are shared with the work group members and revisions to the report are made, if necessary.
- The Technology Assessment Committee reviews the comments and the work group response and makes the final decision regarding approval of the report for distribution.
- Reports are reviewed bi-annually and revised, if warranted.

Description of Technology/Procedure

Osteoporosis is the most common metabolic bone disorder. It is characterized by abnormalities in the amount and architectural arrangement of bone tissue which lead to impaired skeletal strength and an increased susceptibility to fractures (Consensus Development Conference, 1993). A World Health Organization study group recommended that osteoporosis be defined operationally in white women as bone mineral density levels more than 2.5 standard deviations below the young normal mean (Kanis, Melton, Christiansen, Johnston, Khaltaev, 1994). No comparable definition has been devised for nonwhite women or for men. The World Health Organization also defined an intermediate group of women whose bone density was normal but low (more than 1 standard deviation below the young normal mean). This group was labeled "osteopenia," which is somewhat confusing since osteopenia has traditionally been a nonspecific term referring to skeletal demineralization regardless of cause (van Kuijk & Genant, 1994).

While the diagnosis of osteoporosis by the World Health Organization criterion depends on a measurement of bone density, most patients come to clinical attention because of a fracture caused by moderate or minimal trauma. Limb fractures (e.g., wrist, hip) are clinically evident but osteoporotic vertebral fractures may be asymptomatic and first revealed on a routine chest radiograph or they may be associated with chronic mild back pain. Diagnosis of osteoporosis when the fracture is initiated by significant trauma (e.g., motor vehicle accident) requires additional information such densitometry (Kaltenborn, 1992).

Primary osteoporosis can affect any age group. Putative risk factors for primary osteoporosis include smoking, inactivity, delayed puberty, low body weight, excessive alcohol intake, white/Asiatic ethnicity, family history, small body frame, and low calcium intake (Kaltenborn, 1992). Type I (or postmenopausal) primary osteoporosis affects women 15-20 years after the menopause. It is characterized by vertebral crush type fractures, Colles' fractures, and tooth loss (Riggs & Melton, 1986). Vertebral body and radial head fractures are common because of the larger percentage of trabecular (or cancellous) bone in those areas. Trabecular bone is more responsive to estrogen deficiency (Kaltenborn, 1992). Type II (or senile) primary osteoporosis mainly occurs in men and women 70 and older. It is characterized by hip and vertebral wedge type fractures (Riggs & Melton, 1986). Greater cortical bone loss is experienced in this age group.

Secondary osteoporosis can also affect any age group. The causes of secondary osteoporosis include corticosteroid excess, hyperthyroidism, multiple myeloma, malnutrition, estrogen deficiency, hyperparathyroidism, genetic factors, and miscellaneous drugs (Kaltenborn, 1992).

Bone remodeling is an ongoing process; osteoclasts resorb bone and osteoblasts form bone throughout life. Sex, race, nutrition, exercise, and overall health influence the peak bone mass level achieved (Office of Medical Applications Consensus Conference, 1984; Riggs and Melton, 1986). Recent studies have observed that femoral bone mineral density (BMD) peaks by about age 20 years (Matkovic et al., 1994; Looker et al., 1995b). Bone loss is the result of a breakdown in the link between the processes of formation and resorption. Cross-sectional data from Looker et al. (1995b) found the greatest difference in femoral neck BMD between women 40 to 49 years of age and women 50 to 59 years of age. Between age 50 to 59 years and 60 to 69 years (and beyond) the degree of bone loss decreased somewhat but remained elevated above values observed for pre-menopausal age groups. Riggs and Melton (1986) attributed slower phases of bone loss to normal osteoclast activity but reduced osteoblast activity while the accelerated bone loss phases were reported to be related to greater osteoclast activity.

Prevention of osteoporosis is achieved by creating an environment and basic life-style that will facilitate the development and preservation of a high peak bone mass. This includes attention to nutrition, exercise, and other risk factors (van Kuijk & Genant, 1994). Prevention is preferable to treatment (Christiansen, 1994) since there are no effective, safe methods for restoring high quality bone (Consensus Development Conference, 1993). Consequently, there is a need to identify those who are at the greatest risk so that prophylactic therapy may be prescribed (Riggs & Melton, 1986; Cummings & Black, 1986; Davis, 1987; Melton, Eddy, & Johnston, 1990; Johnston, Slemenda, & Melton, 1991; Repa-Eschen, 1993; van Kuijk & Genant, 1994). The benefits from preventative programs for selected populations can offset the costs of the programs (Repa-Eschen, 1993). Although prevention is preferable, many patients exist and will continue to exist who present with osteoporosis. The focus of this report is on identifying those patients and initiating treatment.

This report addresses the issue of densitometry in women. It should not be construed that testing is not appropriate for men, however, given the lack of published data, it would be impossible to comment on the safety and efficacy of densitometry for men.

It is also recognized that bone density is not the only risk factor for bone fracture. However, the relative contributions of other factors (such as bone microarchitecture, propensity to fall, and bone turnover) have yet to be determined.

A variety of methods (ranging from risk factor analyses to histological analyses of bone biopsies) have been proposed to select those patients who are at risk. In screening for osteoporosis, the goal is to predict the likelihood of future fracture (Davis, 1987; Melton et al., 1990; Johnston et al., 1991; van Kuijk & Genant, 1994).

There is debate over the use of central (hip or spine) vs. peripheral (typically radius or calcaneus) site measurements. Osteoporosis is a systemic disease. Therefore, it would be expected that hip and spine fracture risk could be determined based on BMD assessed at a peripheral measurement site. Peripheral sites are technologically easier to measure, less prone to artifacts caused by overlying tissue, and involve a smaller radiation dose at a greater distance from radiation-sensitive organs. The systems are generally lower in cost and many are portable (Glüer, Jergas, & Hans, 1997). However, the role of peripheral measurements of bone density and their relationship with the central measurements at the hip and spine remains controversial (Faulkner, 1998; Miller et al., 1998).

The most widely available, accurate, and precise methods are reviewed below and summarized in Table 1. A discussion of alternative screening procedures is found later in this report.

Single-Photon Absorptiometry

Single-photon absorptiometry (SPA) was the first commercially available technique for measuring bone mineral density (Davis, 1986; Melton et al., 1990; Genant, Faulkner, Glüer, & Engelke, 1993; van Kuijk & Genant, 1994; Erlichman & Holohan, 1995) but it is being replaced by single-energy x-ray absorptiometry (see below). A highly collimated monoenergetic photon beam from a radionuclide source (typically iodine-125) is passed across the measurement site. The transmitted radiation is measured by a detector. Differences in photon absorption between bone and soft tissue permit the calculation of bone mineral content (BMC, grams) or bone mineral density (BMD, g/cm^2). SPA requires a constant soft tissue path length. To achieve this, the measurement site is surrounded by a known thickness of tissue-equivalent material (e.g., a water bath). As a result, this technique is limited to peripheral sites. The most commonly measured site is the distal third of the radius (95 percent cortical bone). Other sites include the distal radius and os calcis (40 and 95 percent trabecular bone, respectively). SPA cannot differentiate between cortical bone and trabecular bone. The accuracy error (defined as the ability of the

measurement to reflect the true bone mass value) is 3 to 8 percent in clinical settings. The precision error (defined as the reproducibility of the value upon repeat measurement) is 2 to 5 percent (also in clinical settings) due largely to errors in repositioning and the low (and constantly decreasing) intensity of the source. The iodine-125 source must be changed several times each year and can therefore adversely affect the reliability of the system. Patient acceptance is very good and scan times are relatively short. Radiation doses are low with negligible whole-body doses.

Single-Energy X-Ray Absorptiometry

Greater precision and speed of scan can be achieved by replacing the SPA radiation source with an x-ray source. Single-energy x-ray absorptiometry (SXA) is also more cost-effective than SPA since the need to replace the iodine-125 source is eliminated (Erlichman & Holohan, 1995; AACE, 1996). It is still necessary to have a tissue-equivalent material surrounding the measurement site (AACE, 1996). Kelly, Crane, and Baran (1994) reported short-term *in vivo* precision errors for the measurement of forearm BMC of 0.66% and BMD of 1.05%. These values were obtained from testing 9 subjects five times on the same day (with repositioning). SXA values for BMC and BMD were highly correlated with dual-energy x-ray absorptiometry (DXA) values (r=0.97 and r=0.96, respectively). Hoshi, Yamada, Tsukikawa, Sugano, Endo, and Sato (1993) noted that the time of the screening was 4 minutes (compared to 10 to 15 minutes for DXA) and that the correlation between vertebral bone mineral density (measured with DXA) and the calcaneus density (measured with SXA) was high (r=0.80).

Dual-Photon Absorptiometry

Dual-photon absorptiometry (DPA) eliminates the need for a constant tissue-equivalent path by emitting photons at two different energies (Davis, 1986; Melton et al., 1990; Genant et al., 1993; van Kuijk & Genant, 1994; Erlichman & Holohan, 1995). The source, typically gadolinium-153 radioisotope, must be replaced every 12 to 18 months. The effective photon energies are typically 44 and 100 keV. DPA can be used to measure the bone mineral density of the lumbar spine, proximal femur, and total body. Like SPA, DPA can measure both BMC and BMD but it cannot distinguish between cortical and trabecular bone. The accuracy and precision error values vary depending on the measurement site. For the lumbar spine, the accuracy error is 3 to 10 percent while the precision error is 2 to 5 percent. At the proximal femur site, the accuracy and precision errors are 3 to 8 percent and 2 to 4 percent, respectively. Patient acceptance is good but scanning times range from 20 minutes for a regional measurement to 60 minutes for a total body measurement. Radiation doses vary depending on the scan site. DPA devices are still being used but are no longer being manufactured.

Dual-Energy X-Ray Absorptiometry

DXA is an upgraded version of DPA with an x-ray source replacing the radioisotope source (Melton et al., 1990; Genant et al., 1993; van Kuijk & Genant, 1994; Erlichman & Holohan, 1995). Dual energy peaks are generated by either filtering or by switching the tube voltage supply. In addition to stability, the x-ray tube allows increased precision, reduced scan times, and reduced radiation doses. DXA systems can be used to measure the BMD of the lumbar spine (anterior-posterior and lateral views), hip, and total body. If a physician suspects that a patient has significant degenerative arthritis, scoliosis, lumbar fractures, a calcified aorta, or other conditions that might affect the lumbar spine bone mineral density reading, the hip site should be considered for testing since these factors rarely affect that site. The accuracy error is similar to DPA (3 to 9 percent) and precision errors are 0.5 to 3 percent.

The implementation and interpretation of DXA testing has continued to evolve. In a comparison of anterior-posterior (AP) and lateral spine measures in healthy, ambulatory

women 65 years of age and older, Greenspan, Maitland-Ramsey, and Myers (1996) found that the mean BMD value for the lateral spine was –3.1 standard deviations from the adult peak BMD (using the manufacturer's normative database) while the corresponding value for the AP spine was –1.7 standard deviations ($p<0.001$). Based on the lateral measures, 66.4% of the women would have been classified as osteoporotic; based on the AP measures, 29.2% would have been classified as osteoporotic (according to the WHO definition). However, these results should be interpreted with caution. Differences in classification between sites may be artifactual because of a lack of standardization of the normative data on which the determination of t-scores is based (Miller et al., 1998). A normative data base for bone density of the hip has been developed as part of the third National Health and Nutrition Examination Survey (NHANES III) and is currently available (Looker et al., 1995a). Similar data for the spine is expected in the near future. In addition, although the study of the AP spine may be confounded by calcification, the coefficient of variation (CV) for the lateral spine is typically higher than the AP spine raising concern about the value of these measures.

Although peripheral sites can be assessed with DXA, modification of the equipment and software is necessary. Recently, a system designed specifically for peripheral sites has been introduced (pDXA). Mole, McMurdo, and Paterson (1998) compared pDXA with SPA of the forearm and DXA of the spine or hip. The correlation between pDXA and SPA (determined with data from 57 people) ranged from $r=0.71$ to $r=0.94$ (depending on the variables studied). The correlation between the pDXA and DXA z-scores ranged from $r=0.29$ to $r=0.74$ (z-scores were used since the same sites could not be assessed). Overall, pDXA was found to have somewhat lower precision than SPA. The advantages of pDXA are a faster scan time and no need to use a water bath.

Another recent development is fan beam DXA. The key feature is a shorter scan time. The disadvantage is lower precision. Patel, Seah, Blake, Jefferies, Crane, and Fogelman (1996) determined the *in vivo* precision of a fan beam system operating with a scan time of 10 seconds (turbo mode). Two scans were done for each of 37 patients (with repositioning between scans). The CV values obtained were 1.3% for the AP spine, 2.0% for the total hip, 2.4% for the trochanter, 2.5% for the femoral neck, 2.6% for the intertrochanter, and 5.6% for the Ward's triangle subregion. The authors also studied the concordance (in 151 patients) between BMD measured with the 10 second scan and BMD measured with scan times of either 30 seconds, 60 seconds, or 2 minutes. The turbo scan BMD values agreed with the normal scan BMD values only for the femoral neck and Ward's triangle regions. Differences of 2.9% and 3.1% were observed for the AP spine and intertrochanteric region of the hip, respectively. Due to the poorer precision, the use of turbo mode scans was not recommended for longitudinal studies.

Franck, Munz, and Scherrer (1997) used a fan beam (FB) system (2 minute scan) and a single beam (SB) system (5 minute scan) to evaluate BMD of the left and right hips. All of the BMD values measured using the FB mode were lower (approximately 2%) than those measured using SB mode ($p<0.008$). Bilateral differences in total BMD were also noted (regardless of scan time).

Quantitative Computed Tomography

Quantitative computed tomography (QCT) can estimate separate bone densities for the trabecular and cortical bone compartments (Davis, 1986; Melton et al., 1990; Genant et al., 1993; van Kuijk & Genant, 1994; Erlichman & Holohan, 1995). It is the only technique that can measure a true three-dimensional density. Like the other procedures, QCT is based on differential absorption of ionizing radiation by calcified tissues. Measurements are compared to a standard reference. Initially, a dipotassium hydrogen phosphate liquid solution was used but more recently, solid calibration standards based on calcium hydroxyapatite have been used. Single- and dual-energy procedures are available. Variations in the fat content within the marrow can affect the accuracy of QCT by as much

as 20 to 30 percent (Alhava, 1991; Baran, 1994). Accuracy is improved with the dual-energy system but increased radiation doses are a limitation. Dual-energy QCT is recommended for research purposes only. The accuracy error for single-energy QCT is 5 to 15 percent while the precision error is 2 to 6 percent. Better precision is obtained with dedicated (rather than multipurpose) scanners. QCT has generally been used to measure the lumbar spine. However, measurements of the radius are also possible using a special purpose CT system (Rüeggsegger, Durand, & Dambacher, 1991). The precision and accuracy errors of peripheral QCT of the radius are both reported to be 0.5 to 1 percent (AACE, 1995). This procedure requires less time, offers improved precision and accuracy, and exposes the patient to lower levels of radiation. Trabecular bone in the femoral neck has also been examined, however, the complexity of the anatomy may make reproducible measurements difficult (Erlichman & Holohan, 1995). In addition, measurement of just trabecular bone may not be adequate to predict femoral fracture risk (Sartoris, André, Resnick, & Resnick, 1986). In general, patient acceptance of QCT is good with a scan time of 15 to 20 minutes. Radiation doses (even for single-energy QCT) are higher than for any of the other techniques.

Radiographic Absorptiometry

Radiographic absorptiometry (RA), formerly referred to as photodensitometry, is a radiologic technique primarily used to evaluate the small bones of the hand (Erlichman & Holohan, 1995; Yates, Ross, Lydick, & Epstein, 1995). To compensate for variability due to voltage setting, exposure time, film quality, and film processing, an aluminum alloy reference wedge is placed in the radiographic field (Cosman, Herrington, Himmelstein, & Lindsay, 1991; Yates et al., 1995). The analysis is no longer labor intensive and operator dependent as computerized systems are used to capture and digitize the images. Cosman et al. (1991) observed correlations of 0.58 to 0.78 between density determined by RA and standard densities (SPA and DPA). The accuracy error is 5 to 10 (correcting for soft tissue and fat overlying the bones is a problem). The precision error is between 2 and 4 percent (Erlichman & Holohan, 1995). Cosman et al. (1991) concluded that RA might be useful as a screening tool in settings where DPA or DXA are impossible. In Minnesota, Chapter 47.30.1210 of the Rules for Ionizing Radiation from the Minnesota Department of Health prohibits the use of non-screened film in cardboard holders. This type of holder requires radiation doses that are 6 to 10 times higher than with screened, cassette film holders. The cardboard holders would be used when the film must be sent out for processing and interpretation.

Ultrasound

Quantitative ultrasound (QUS) can provide information on bone strength and microarchitecture (Genant et al., 1996; van Daele, Burger, De Laet, & Pols, 1996; Njeh, Boivin, & Langton, 1997; Hayes, 1998) that might be related to BMD and fracture risk. It has been suggested that since bone BMD explains only 70-75% of the variance in bone strength, the remainder is likely due to other factors such as architecture, porosity, and connectivity (Glüer, 1997; Njeh et al., 1997).

The advantages of QUS are that it is non-invasive, involves no exposure to ionizing radiation, is less expensive than x-ray based systems, it can be portable, and that, because of the absence of ionizing radiation, there are fewer regulations regarding the personnel and facilities required for operation of the system (Johnston & Melton, 1995; Ross et al., 1995; Glüer, 1997). The typical measures are velocity (time for the ultrasound wave to travel from the transmitting transducer to the receiving transducer often expressed as the speed of sound [SOS]) and attenuation (a loss of energy reflecting both scattering and absorption often expressed as broadband ultrasound attenuation [BUA]) (van Daele et al., 1996; Glüer et al., 1997; Njeh et al., 1997; Hayes, 1998). It has been suggested that the two measures be combined into a single term. One group has labeled this "stiffness" (not the

same as true mechanical stiffness) and another has labeled this the "quantitative ultrasound index" (QUI) (Glüer et al., 1997). Several QUS systems have been approved by the FDA.

An International Quantitative Ultrasound Consensus Group (Glüer, 1997), concluded that QUS was promising. They suggested that a combination of elasticity, structure and density information might provide a more sensitive indicator of fracture risk than density alone. There is a need for normative data (including data from men and younger women), quality assurance programs, standardization, and attention to precision, sensitivity, and accuracy. For the diagnosis of osteoporosis, the group found there was no agreement on how to interpret the QUS results. They suggested using BMD in parallel with QUS. For assessment of fracture risk, the only prospective data is from use of a water-based calcaneal QUS system (see below). Validation of other devices is needed. For monitoring of skeletal changes, QUS is not recommended at present. As a result of the somewhat higher precision errors, the time between testing would be greater than that recommended for BMD. It was also suggested that QUS might be used to select patients for referral for BMD (Glüer et al., 1997).

Table 1. Comparison of techniques for measuring bone mass (Adapted from Kelly et al., 1994; Erlichman & Holohan, 1995; AACE, 1996)

Technique	Equipment Cost ($)	Scan Charges ($)	Site; Scan Time (min)	Precision Error (%)	Accuracy Error (%)	Radiation Exposure (mrem)
SPA	20,000-30,000	50-150	Radius, calcaneus; 5-15	2-5	3-8	2-5
SXA	NR[a]	NR	Radius, calcaneus	0.7-1.1	NR	NR
DPA	30,000-65,000	150-300	Spine, hip; 20-40	2-5	3-10	5-10
DXA	60,000-100,000	185 1 site 225 2 sites	Spine, hip; 5-10	0.5-3	3-9	<5
QCT	5,000-15,000 (adapt existing CT scanner)	150-400	Spine; 10-30	2-6	5-15	100-1,000
RA	Existing x-ray equipment	90-160	Hand; 5-10	2-4	5-10	10-100
Ultrasound	25,000	45	Peripheral skeleton	4	NR	0

[a]NR=Not Reported

Precision and Accuracy

The precision and accuracy error values of each method should be considered in interpreting the results of testing (Hassager, Jensen, Godfredsen, & Christiansen, 1991; Alexeeva et al., 1994; Christiansen, 1994; Erlichman & Holohan, 1995). Precision error is the variability occurring with repeated measurement of the same stable object. It is usually expressed as a coefficient of variation [(standard deviation/mean)*100]. Precision error is important if repeated measurements are to be made on a single patient and in longitudinal prospective studies of groups of patients. Differences in the positioning of the body or segment at repeat testings, changes in soft tissue composition over time, and

differences in the strength of the radionuclide source (for SPA and DPA) can affect the precision error (Erlichman & Holohan, 1995). It is important to note that a given technique will likely have different precision error values depending on the site being measured (e.g., DXA of the spine or hip) and the test center. Each test center needs to determine its precision error values based on the equipment and procedures used (Alexeeva et al., 1994).

The precision errors must be evaluated with respect to the rate of bone loss (Erlichman & Holohan, 1995). Rate of bone loss will vary depending on the measurement site and the patient's age or menopausal status. For premenopausal women (ages 22 to 54 years), a mean decrease in BMD of the radius (determined by SPA) of 5.6% over a 5 year period was reported by Sowers, Clark, Hollis, Wallace, and Jannausch (1992). In the immediate postmenopausal period, bone loss values of 2.8% per year (proximal forearm) to 4.4% per year (spine) were reported using DPA as the measurement tool. The differences between sites are related to the different compositions of bone (cortical or trabecular) at those sites (Alexeeva et al., 1994). Black (1995) reported typical bone loss values of 1.0% per year or less (higher values more likely for the radius, calcaneus, hip with lower values for the spine) for women up to age 70 to 80 years. Recent evidence suggests that bone loss may accelerate in older women.

To reliably (with 95 percent confidence) detect a reduction in bone mass of 1 to 3 percent given precision errors of 1 to 6 percent, the time interval between measurements would range from 1 to 17 years (assuming that accuracy was constant). Thus, for a densitometry system with a 2 percent precision error, if the annual bone loss was 1 percent, the minimum interval between densitometry measures (in order to document bone loss) would be 5.5 years. In a clinical setting, the precision errors are likely higher than the 2 percent value used in this example. Table 2 summarizes these calculations.

A follow-up interval of 4 to 5 years may be adequate for assessing rate of bone loss (in most cases) but would not be useful for monitoring the effects of treatment. Given the limitations of the measurement systems, yearly densitometry is not clinically indicated (Erlichman & Holohan, 1995).

Table 2. Approximate follow-up time for the reliable detection of bone mass loss (from Erlichman & Holohan, 1995)

Precision Error (%)	Bone Loss (%) (annual)	Differences in Measurement (%)[a]	Approximate follow-up time (years)
1	1	2.8	2.8
1	3	2.8	0.9
2	1	5.5	5.5
2	3	5.5	1.9
3	1	8.3	8.3
3	3	8.3	2.8
4	1	11.1	11.1
4	3	11.1	3.7
5	1	13.3	13.3
5	3	13.3	4.4
6	1	16.6	16.6
6	3	16.6	5.5

[a]Two measurements would have to differ by more than this amount to be 95% confident that a real change had occurred

Accuracy error is the total error made in estimating the true value (Hassager et al., 1991; Alexeeva et al., 1994; Christiansen, 1994; Erlichman & Holohan, 1995). Cadaver studies are needed to determine the accuracy error. Sources of inaccuracy include the irregular shape of the vertebrae, source life (SPA and DPA), osteophytes and soft tissue calcification, and variable densities of soft tissues (Alexeeva et al., 1994). Accuracy error must be evaluated with respect to the biological variation of bone mass in the population. Biological variation refers to the variance in measurement of bone mass in the population of interest and is expressed as a coefficient of variation. If the accuracy error is small compared with the biological variation, the diagnosis will be reliable. If the accuracy error is large compared with the biological variation, the diagnosis will be uncertain. In healthy women just after the menopause, the biological variation of bone mass is about 15 percent. Spinal bone mass measurement has an accuracy error of up to 10 percent (Hassager et al., 1991; Christiansen, 1994).

Potential Uses

It is generally accepted that a screening test that could accurately assess the risk for future hip fractures would be of great value (Cummings & Black, 1986; Ott, 1986; Davis, 1987; Melton et al., 1990; Johnston et al., 1991; van Kuijk & Genant, 1994). However, given that bone density is a continuous variable, there is no true "fracture threshold." Thus, the choice of an intervention threshold is somewhat arbitrary resulting in the potential for misclassification (Alexeeva et al., 1994). It has been suggested that until a specific protocol has been developed for general screening of some portion of the population, mass screening cannot be justified (Johnston & Melton, 1995).

Densitometry screening measurements would be indicated for those individuals who are at risk for osteoporosis and who are willing to accept intervention therapies (Riggs & Melton, 1986; Repa-Eschen, 1993; Comptson, Cooper, and Kanis, 1995; Johnston & Melton, 1995; Ribot, Tremollieres, & Pouilles, 1995). Determination of risk could be based on the historical risk factors (Table 3).

Table 3. Major historical risk factors for osteoporosis in women (from Riggs & Melton, 1986)

Postmenopausal (within 20 years after menopause)	Nulliparity
White or Asian	Gastric or small-bowel resection
Premature menopause	Long-term use of glucocorticoid therapy
Positive family history	Long-term use of anticonvulsants
Short stature and small bones	Hyperparathyroidism
Leanness	Thyrotoxicosis
Low calcium intake	Smoking
Inactivity	Heavy alcohol use

The American Association of Clinical Endocrinologists (AACE) guidelines (AACE, 1996) recommend that BMD measurements should be performed in the following settings:

1. For risk assessment in perimenopausal or postmenopausal women who are concerned about osteoporosis and willing to accept available interventions;
2. In women with x-ray findings that suggest the presence of osteoporosis;

3. In women beginning or receiving long term glucocorticoid therapy, provided intervention is an option;
4. For perimenopausal or postmenopausal women with asymptomatic primary hyperparathyroidism in whom evidence of skeletal loss may result in parathyroidectomy;
5. For women undergoing treatment for osteoporosis, as a tool for monitoring therapeutic response.

Perimenopausal women should be counseled on the benefits of estrogen replacement therapy as should women who are estrogen deficient as a result of premature menopause, bulimia, anorexia nervosa, amenorrhea, or some other condition. Women who are on long term estrogen replacement for some other reason may be receiving an adequate treatment for osteoporosis and *generally* do not need a bone density measurement (AACE, 1996). However, in light of observed cases of bone mass loss despite long term, well-tolerated hormone replacement therapy (HRT), and the fact that many women do not begin HRT until substantial bone loss has occurred, testing may be appropriate for this group (see National Osteoporosis Foundation recommendations below).

The National Osteoporosis Foundation (National Osteoporosis Foundation, 1998b) has developed nomograms to assist physicians in determining whether diagnostic testing and/or treatment are appropriate. The nomograms take into consideration age, risk factors, prior fractures, and prior treatment. They recommend that the decision to test be made in the context of the particular treatment being contemplated. Some treatments may not require testing.

In the accompanying physician's guide (National Osteoporosis Foundation, 1998a), BMD testing was recommended for the following groups. *The recommendations were prefaced by a statement that the decision to test for BMD should be based on an individual's risk profile and that testing is never indicated unless the results could influence a treatment decision.*

1. All postmenopausal women under age 65 who have one or more additional risk factors for osteoporosis (besides menopause);
2. All women aged 65 or older regardless of additional risk factors;
3. Postmenopausal women who present with fractures (to confirm diagnosis and determine disease severity);
4. Women who are considering therapy for osteoporosis, if BMD testing would facilitate the decision; and
5. Women who have been on HRT for prolonged periods.

DXA, SXA, pDXA, RA, QCT, and QUS were recommended recognizing that QUS measurements are less precise than DXA or SXA.

Medicare coverage of and payment for bone mass measurements was recently outlined (Department of Health and Human Services, 1998). Effective July 1, 1998, bone mass measurement is covered for "medically necessary approved measurements" performed for a "qualified individual." A qualified individual would fall into at least one of following diagnostic categories:

1. Estrogen deficient women at clinical risk for osteoporosis (based on medical history or other findings);
2. An individual with vertebral abnormalities (as demonstrated by x-ray to be indicative of osteoporosis, low bone mass, or vertebral fracture);
3. An individual receiving long-term glucocorticoid (steroid) therapy (equivalent to 7.5 mg of prednisone, or greater, per day for more than 3 months or if the expected duration of such therapy is more than 3 months);
4. An individual with primary hyperparathyroidism; and/or
5. An individual being monitored to assess the response to, or efficacy of, an approved osteoporosis drug therapy.

The measurements may be made with "a radiological, radioisotopic, or other procedure approved by the FDA." Coverage for follow-up bone mass measurements is limited to one measurement every 2 years unless more frequent measurement is medically necessary. Thus, patients receiving long-term glucocorticoid therapy as described in number 3, above, would be eligible for more frequent testing (6 month intervals are recommended for these patients). It would also allow patients to have a confirmatory baseline bone mass measurement if future measurements are to be made with a different technique than used for the original testing. Biochemical markers, although approved by the FDA, are not included in this benefit although they may be covered under the clinical laboratory fee schedule.

The American College of Rheumatology (1996) guidelines recommend that when glucocorticoid therapy is initiated with doses of 7.5 mg of prednisone per day, baseline BMD measurements should be done with repeat testing in 6 to 12 months. These guidelines are currently being reviewed.

Other potential indications for bone mass measurement have been identified but have not been substantiated by scientific data (Scientific Advisory Board, 1989). These indicators might signal the need for bone mass measurement for the purpose of medical evaluation of an individual patient but are not relevant for general screening purposes.

Contraindications

Bone mass measurements are not recommended for the following groups (Scientific Advisory Board, 1989):

1. Women who were to be put on long-term estrogen replacement therapy for other reasons (although BMD testing would be useful to guide treatment modifications if the patient presented with insufficiency fractures);
2. Women for whom estrogen is contraindicated and who are unwilling to consider an alternative therapy;
3. Women who refuse to take estrogen or participate in some other therapy to slow bone loss;
4. Women with vertebral abnormalities whose treatment would not be altered by the outcome of the measurement or in those whose radiographic evidence is unequivocal;
5. Patients who are receiving glucocorticoid treatment if no alterations in their therapy were clinically possible;
6. Women with primary hyperparathyroidism who have been identified as candidates for surgery on some other basis or if they would refuse to have surgery or other medical therapy.

In general, screening measurements are not indicated unless they will influence patient or physician behavior. There must be an organized plan for patient management and referral depending on the outcome of the measurement. Screening is also contraindicated when standardization and quality control procedures are insufficient to insure accurate results (Scientific Advisory Board, 1989).

Densitometry is reported to be of limited use for women who are older than 70 years of age. Age-related changes in calcium absorption, parathyroid hormone levels, and vitamin D levels result in a thinning of both cortical and trabecular bone (despite a decline in the rate of bone loss compared to the postmenopausal period). Most elderly patients have bone density values that fall below the fracture threshold as determined through densitometry. It is also more difficult to measure vertebral bone density because of the high degree of degenerative joint disease, vascular calcifications, compression fractures, and an increase in marrow fat (Resnick & Greenspan, 1989). By age 70, most white women

will have bone mass measures that are within the range observed in those with hip fractures (Cummings & Black, 1986). However, the depth and severity of osteoporosis at age 70 and above may influence the choice of therapy; some would recommend screening for this age group if the documentation would lead to therapeutic intervention.

In addition, recent evidence has found that bone loss continues past the age of 80 years (Black, 1995). Garnero, Shih, Gineyts, Karpf, and Delmas (1994) found a significantly higher rate of bone turnover in late postmenopausal (by a minimum of 5 years and an average of 16 years) women than in premenopausal women. They suggested that bone turnover remains elevated long after menopause. Elevated bone resorption (rather than decreased bone formation) may result in continued, age-related bone loss. Black (1995) proposed that all women be screened at age 50. Those in the highest risk group based on bone mineral density values and other risk factors (1 to 2 percent of the population) would be treated immediately. Those in the moderate risk group (the next 15 to 20 percent) would receive dietary calcium and be screened again in 5 years. The remaining 80 percent would not undergo a second screening until age 65 to 70 unless an intervening risk factor developed. Screening at age 65 would provide a better estimate of risk of fracture after age 80 (when the majority of fractures occur). It is possible that starting estrogen therapy at age 65 would result in only slightly lower bone mass at age 80 than if therapy was begun at age 50. This would be beneficial both in terms of cost and of minimizing the exposure to risks associated with long-term therapy.

Efficacy of Treatment or Procedure

PREDICTION OF FRACTURES

The efficacy of screening for bone density is related to the potential to provide information that will identify individuals who are at risk for fractures, especially hip fractures. It is recognized that factors other than bone mass may be important in determining who is more likely to fracture their hip. These factors would include poor eyesight, proneness to falling, ability to protect oneself in the event of a fall, geometric structure of the bones, and bone quality (Cummings & Black, 1986; Cummings et al., 1990; Melton et al., 1990; Kaltenborn, 1992).

Randomized, controlled trials to determine whether screening is effective in reducing the number of fractures are not feasible since the latent period between the onset of bone loss and the occurrence of hip fractures may be 20 to 30 years (Melton et al., 1990).

A review of retrospective studies (where bone mineral density was evaluated at the time of fracture and compared to age-matched controls) found considerable overlap between subjects with and without fractures (Law, Wald, & Meade, 1991). On this basis, the authors concluded that there was no scientific case for screening. However, it is inappropriate to use bone mass measurements as diagnostic tests for fractures (Davis, Vogel, Ross, & Wasnich, 1989). BMC is a risk factor with a continuous distribution, and risk factor analysis techniques can be used to determine the distribution of risks in the population, to calculate the relative risks (risk ratios), and to determine the percentage of fractures attributable to low values of BMC.

The most useful data is obtained from prospective studies where bone mass is measured and the subjects followed, keeping records of fractures (Johnston & Melton, 1995). Many prospective studies have been published all of which consistently show that bone density measurements can predict the risk of fractures; a few representative studies are summarized below.

Hui, Slemenda, & Johnston (1989) followed 521 white women for up to 15 years. The subjects included 356 free-living women and 135 retirement home residents ranging in age from less than 40 to greater than 80 years. All subjects were free of disease and not receiving any medication known to affect bone metabolism. At their initial visit, bone density was evaluated using SPA at the midshaft of the radius. Vertebral fractures and automobile related fractures were excluded from the analysis. Average follow-up times were 6.7 years for the free-living group and 5.5 years for the retirement home group. During the follow-up, 89 subjects (35 free-living and 54 retirement home) had a total of 138 fractures. For every 0.1 g/cm decrement in bone mass (approximately 1 standard deviation), the relative risk of fracture at any site was 2.2 for the free-living and 1.5 for the retirement home residents.

A group of 9703 nonblack women age 65 and older were studied by Cummings et al. (1990). SPA was used to measure BMC and BMD in the distal radius, the proximal radius, and the calcaneus. The patients were followed for an average of 1.6 years. During that time, 53 participants experienced a first fracture of the proximal femur (48 fractures resulted from a fall and 5 were spontaneous, i.e., not the result of a fall or trauma). The age-adjusted relative risks of a hip fracture for a decrease of 1 standard deviation in BMD were 1.66 for the calcaneus, 1.55 for the distal radius, and 1.41 for the proximal radius. Similar relationships were found for BMC. None of the measurements was superior for predicting hip fractures. However, even when adjusting for bone mineral density, age remained a significant predictor of hip fracture with the risk of hip fracture doubling for each 10 year increase in age. It was concluded that decreased bone density was just one factor in the age-related increased risk of hip fracture in older women.

BMC and BMD of the hip and lumbar spine (using DXA) were determined on 8134 of these patients approximately 2 years after the initial measurements (Cummings et al., 1993). Follow-up after the second set of measurements averaged 1.8 years. During that time there were 65 hip fractures. Bone density values from all regions of the hip were strongly related to the risk of hip fracture. After adjusting for age, each standard deviation decrease in bone density of the proximal femur increased the fracture risk from 2.5 to 2.8 times (depending on the exact site of the measurement). The relation between bone mineral density and risk of hip fracture was stronger than that between bone mineral content and risk of fracture. The authors concluded that measurements of bone density in the proximal femur provide a better prediction of hip fractures than measurements at other skeletal sites.

Gärdsell, Johnell, and Nilsson (1991) sought to determine whether initially low bone mass or subsequent bone loss was the most important predictor of future fragility fractures. BMC of the middle and distal radius was determined (by SPA) for a total of 1076 women. After an average of 14.6 years, 366 of these patients agreed to a repeat test using the same equipment and techniques. All fractures during the intervening years were also recorded. The initial BMC was less (regardless of age group) in those women who had a fragility fracture during the years between the BMC evaluations. Rate of bone loss was positively correlated with bone mass such that the greater the bone mass, the greater the loss, however, rate of loss did not differ between fracture and non-fracture women making initial bone mass the better predictor. For these subject, peak bone mass occurred before age 40.

Further follow-up data on this group of patients was presented by Düppe, Gärdsell, Nilsson, and Johnell (1997). For the 410 women who were alive in 1994, all fragility fractures (of the vertebrae, proximal humerus, proximal femur, and distal radius), verified by x-ray, were noted. There were a total of 213 fractures (63 vertebrae, 30 humerus, 43 femur, and 77 distal radius). The women were divided into three overlapping age groups based on age at time of BMD testing: ≥40 but <70 years, ≥30 but <50 years, and

≥50 years. The relative risk (adjusted for age) for fracture with a 1 standard deviation reduction in BMD of the proximal femur was determined at 5, 10, 15, 20, and 25 years from the time of measurement. In the ≥40 but <70 years group, the relative risk of hip fracture was 1.66 (95%CI: 1.13-2.46) at 25 years. For the ≥50 years group, the corresponding value was 1.52 (95%CI: 1.02-2.28). For vertebral fractures, the relative risk values were 1.79 (95%CI: 1.22-2.62) for the ≥40 but <70 years group and 1.59 (95%CI: 1.10-2.30) for the ≥50 years group. For all fragility fractures, the relative risk values were 1.33 (95%CI: 1.2-1.73) and 1.20 (95%CI: 0.98-1.69), respectively. These results indicate that a single measure of BMD has a predictive ability for fragility fractures for as long as 25 years.

Wasnich (1993) presented a summary of several prospective studies from the Hawaii Osteoporosis Center. The subjects were all postmenopausal women. In one study, the rate of new vertebral fractures increased 2.0 to 2.4 times for each standard deviation decrease in baseline bone density (including measures of the distal and proximal radius, calcaneus, and lumbar spine). In another study, 423 postmenopausal women with one to four existing spine fractures were evaluated over an average of 2.9 years. Hip bone density was measured with DPA while the spine was evaluated with both DPA and QCT. The increase in the rate of new vertebral fractures per standard deviation decrease in bone density was identical to that of the first study. In both studies, density measurements of more peripheral locations predicted the risk of spine fractures as well as measurements of the spine. Fracture risk (for the spine) was found to increase with decreasing levels of bone density. Women with bone density at or below the mean have a greater risk for fracture than those above the mean.

To assess the relative ability of measurements of bone mineral from different sites to predict fractures, 304 subjects (age range from 30 to 94) were evaluated and then followed for up to 10 years (median 8.3 years) (Melton et al., 1993). DPA was used to determine BMD at the lumbar spine and cervical and intertrochanteric regions of the right proximal femur. SPA was used to measure BMC of the midradius and distal radius. The sites were chosen to represent a range of proportions of cancellous and compact bone. The occurrence of new fractures during the follow-up period was recorded. Ninety-three subjects experienced a total of 163 incident fractures with just over half involving the proximal femur, vertebrae, or distal forearm. Included in this total were 34 fractures resulting from severe trauma and 4 from pathological conditions. Based on 37 subjects with a new vertebral fracture, a 1 standard deviation in lumbar spine BMD was associated with a 1.9-fold increase in the risk of a vertebral fracture. For the other measurement sites, the age adjusted relative risks ranged from 1.3 (distal radius) to 2.6 (midradius). The age adjusted relative risk of hip fractures ranged from 1.3 (midradius) to 2.6 (distal radius). The risk of hip fracture increased 2.4-fold and 2.3-fold, respectively, per standard deviation decrease in femoral neck and trochanter BMD. Age was not related to risk of distal forearm fractures. Per standard deviation decrease in BMD, relative risk increased from 1.5-fold (trochanter and femoral neck BMD) to 2.7-fold (distal radius BMD). The authors concluded that bone mineral measurements from a variety of skeletal sites provide data that will enable the prediction (for at least 8 to 10 years) of moderate trauma fractures such as those that might be associated with osteoporosis. Bone mineral measurements at any of the typically measured sites can be used to group patients according to their overall fracture risk.

Hans et al. (1996) performed baseline measures of calcaneal QUS and femoral neck BMD (DXA) in 5,662 women who were at least 75 years old. The women were then contacted every 4 months to ascertain whether they had sustained a hip fracture. Reported fractures were subsequently confirmed with medical records. Compliance with the follow-up was 97.7%; 2.2% (n=166) refused to continue and 0.1% (n=10) were lost to follow-up. In a mean follow-up time of 2 years, there were 248 deaths (unrelated to hip fracture) and 115 first hip fractures (as a result of minor trauma). Low values of BUA, SOS, and BMD were

associated with increased risk of hip fracture (after adjustment for age). The relative risk associated with a decrease of one standard deviation in BUA was 2.0 (95%CI: 1.6-2.4). For SOS the relative risk was 1.7 (95%CI: 1.4-2.1) and for BMD the relative risk was 1.9 (95%CI: 1.6-2.4). The relative risk values were basically unchanged after controlling for either BMD, BUA, and SOS indicating that the three measures were independent predictors of fracture. If both BMD and BUA were above their respective median values, the rate of fracture was 2.7 per 1000 women. If both BMD and BUA were below their median values, the rate of fracture was 19.6 per 1000 women.

Bauer et al. (1997) studied women over 65 years of age. The 6,189 participants represented about 72% of the original Study of Osteoporotic Fractures cohort known to be alive at 5 years after initial enrollment. Each subject was evaluated for BUA of the right heel and BMD (DXA) of the femoral neck and right calcaneus. Short-term reproducibility of BUA was assessed in 40 unselected subjects. CV values of 4.1% to 5.6% were obtained across the 4 participating clinics. The mean CV was 5.0%. The CV for BMD of the femoral neck was 1.2% (based on data from 2 staff members who visited each of the 4 clinics). The CV for repeated testing of a femoral neck phantom was 0.8%. As in the study by Hans et al. (1996), the women were contacted every 4 months. The mean follow-up was 2 years with a 99% complete follow-up. Hip and other non-spine fractures were recorded. Spine fractures were excluded from the analysis because of the poor reliability of self-reported fractures at that site. All fractures were verified by medical records or radiographs. Hip fractures were subdivided into femoral neck or intertrochanteric fractures. A total of 350 non-spine fractures were reported including 54 hip, 70 wrist, 40 humerus, and 36 ankle fractures. Low BUA and low BMD were similarly associated with an increased risk for hip and non-spine fractures. For BUA, the relative risk for a non-spine fracture was 1.3 (95%CI: 1.2-1.5). For calcaneal BMD, the relative risk for a non-spine fracture was 1.4 (95%CI: 1.2-1.6) while for femoral neck BMD, the relative risk was 1.3 (95%CI: 1.1-1.5). For any hip fracture, the corresponding relative risks (all $p<0.05$) were 2.0 for BUA, 2.2 for calcaneal BMD, and 2.6 for femoral neck BMD. For the 28 trochanteric hip fractures, the relative risk values (all $p<0.05$) were 3.3 for BUA, 3.4 for calcaneal BMD, and 3.9 for femoral neck BMD. For the 26 femoral neck fractures, only the relative risk based on femoral neck BMD was significant (RR=2.0; 95%CI 1.2-3.1). Adjustment for hip BMD did not change the association between low BUA and increased fracture risk. A combination of the BUA and BMD values offered little clinical advantage over femoral neck BMD alone in predicting risk of hip fracture.

A subset of 560 postmenopausal women from the Hawaii Osteoporosis Study were the subjects of the study by Huang et al. (1998). The women underwent baseline bone mass measurements (radiographic absorptiometry and calcaneal QUS), spine radiographs at baseline and at follow-up (a mean of 2.7 years), and follow-up assessment of non-spine fractures between the two examination times. Fractures were identified by self-report with verification by medical records. Only non-violent fractures were studied. The mean age of the subjects at baseline was 74 years (range from 55 to 92 years). Calcaneal BUA was positively associated with fracture risk. For a one standard deviation decrease in BUA, the odds ratios and 95% confidence intervals (adjusted for age) for vertebral fractures, non-spine fractures, and any fractures were 1.50 (1.05-2.16), 1.89 (1.27-2.88), and 1.72 (1.30-2.31), respectively.

Thompson, Taylor, Oliver, and Fisher (1998) reported the relationship between baseline calcaneal QUS and subsequent fractures over a mean follow-up time of 31 months. Fracture data were obtained from practice records of 3,143 women. Thirty seven women were lost to follow-up. Overall, there were 150 fractures (including 2 fractures in 7 women). Baseline data from those who experienced fractures and those who did not indicated significant ($p<0.001$) differences in age (higher among those with fractures), BUA, SOS, and a stiffness index (a combination of BUA and SOS) (all lower among those with fractures). Odds ratios and 95% confidence intervals (adjusted for age) for a one standard deviation decrease in the QUS parameters were determined for all fractures (150 in the study), wrist

fractures (63 in the study), osteoporosis-related fractures (defined as fractures of the proximal humerus, hip, pelvis, and vertebrae, 26 in the study), and other fractures (61 in the study). For all fractures, the values were 1.4 (1.1-1.6), 1.3 (1.1-1.5), and 1.5 (1.3-1.8) for BUA, SOS, and stiffness, respectively. For wrist fractures, the corresponding values were 1.6 (1.2-2.1), 1.5 (1.2-21.), and 1.8 (1.4-2.4). For osteoporosis-related fractures, the values were 1.9 (1.3-2.9), 1.5 (1.0-2.5), and 2.2 (1.4-3.4). For other fractures, the values were 1.0 (0.8-1.3), 1.1 (0.8-1.3), and 1.1 (0.9-1.5). When the data were expressed in terms of age groups (46-55 years, 56-65 years, and 66-75 years), wider confidence intervals (many of which included 1.0) were observed for the youngest age group. The authors recommended caution in using the results from this age group and attributed the findings to the small number of fractures in these women.

It is important to note that the QUS studies cited above used water-based QUS systems. It is unknown whether similar results would have been obtained with dry systems.

SCREENING TESTS

A previous ICSI Technology Assessment Report (1993) identified the criteria by which a screening procedure should be evaluated. The traditional definition of screening was used. A successful screening test would be able to detect the target condition earlier than without screening and would lead to a better clinical outcome for those who are detected early.

The following criteria were identified:

1. Early diagnosis must be scientifically proven to lead to improved clinical outcome.
2. Additional resources are available to confirm the diagnosis and provide care for those who screen positive.
3. Patients who are screened must be willing to comply with subsequent treatment recommendations.
4. The burden of disability from the target disease warrants action.
5. The effectiveness of the individual components of the screening test have been demonstrated prior to their contribution.
6. The cost, accuracy, and acceptability of the screening test are adequate for the population to be screened.

These (or similar) criteria have been applied to the topic of screening for osteoporosis (Cummings & Black, 1986; Alexeeva et al., 1994). Alexeeva et al. (1994) grouped the criteria under four headings: The Disease, The Test, The Intervention, and The Program.

The Disease

The disease for which a screening test is proposed must be an important public health problem and its natural history must be adequately understood (Alexeeva et al., 1994). Data on the number of fractures, mortality, changes in life-style following fracture, and cost support the identification of osteoporosis as a public health issue (Cummings & Black, 1986). The natural history of the disease is well delineated as patterns of change in bone density with age are generally understood. The relationship between bone density and fracture risk has been demonstrated in numerous case-control and prospective studies (Alexeeva et al., 1994).

The Test

The screening test must be simple, safe, acceptable to the population, effective, sensitive, and specific (Alexeeva et al., 1994). The commonly used techniques (discussed earlier in this paper) are all relatively simple and safe with minimal exposure to radiation (Cummings & Black, 1986; Alexeeva et al, 1994). Patient acceptance is good (Rubin & Cummings, 1992) although the acceptability of the tests has not been evaluated in the

context of a mass screening program (Alexeeva et al., 1994). Bone density measurement provides the best predictor of future fracture risk (Alexeeva et al., 1994; Johnston & Melton, 1995). In terms of lifetime risk of fracture (rather than the presence of a fracture), sensitivity would be defined as "the proportion of individuals who would in their lifetime sustain a fracture with a bone mineral density *below* a defined cut-off value." Specificity would be defined as "the proportion of subjects who would not sustain a fracture in their lifetimes with bone mineral density values *above* the cut-off." The sensitivity of a single measurement of bone mass for the prediction of any osteoporotic fracture has been estimated to be between 29 and 80 percent. The specificity of that same prediction has been estimated at 78 to 100 percent. The low values for sensitivity suggest that a substantial proportion of fractures will occur in women who lie in the lower risk groups (Alexeeva et al., 1994). The sensitivity and specificity analysis is confounded by the presence of factors other than bone mineral density that affect fracture risk, however, that does not imply that bone mineral density is not an important risk factor.

The Intervention

Following the screening, there must be an intervention that is accepted and effective. The policy on whom to treat must be agreed upon (Alexeeva, et al., 1994). Therapies for osteoporosis focus on either inhibiting bone resorption or stimulating bone formation (Consensus Development Conference, 1991; Riggs & Melton, 1992). It is also important to attempt to reduce the risk of falling (Alexeeva et al., 1994).

Inhibitors of bone resorption include estrogen, calcitonin, bisphosphonates, and calcium. Estrogen slows the rate of bone loss following the menopause. It also produces increases in bone mass in women with established osteoporosis (Riggs & Melton, 1992) and may be effectively introduced (or reintroduced) in women age 65 to 70 (Consensus Development Conference, 1991; Bergman, 1996). However, estrogen has side effects that must be considered in determining whether HRT is appropriate for an individual.

Kohrt and Birge (1995) evaluated the effects of 1 year of HRT (estrogen and progestogen) in 12 women who were 10 or more years past the menopause. A control group of 12 women, matched with respect to body weight and initial BMD of the lumbar spine and femoral neck, received no therapy. Significant increases were observed in the BMD of the total body (1.4 percent), lumbar spine (5.0 percent), femoral neck (3.2 percent), trochanter (3.2 percent) and Ward's triangle (6.6 percent). A slight decrease in BMD was seen at the ultradistal radius.

In the case-control study presented by Michaëlsson et al. (1998), use of HRT was evaluated in women with hip fractures (n=1327) and women without hip fractures (n=3262). Participants in the study completed a written questionnaire. Use of HRT at any time was associated with a reduction in risk of hip fracture (OR=0.58; 95%CI: 0.46-0.75). For current users, the odds ratio was 0.36 (95%CI: 0.24-0.53). The odds ratios for combined estrogen-progestin protocols were lower than for estrogens alone. Among long-duration users of HRT, the reduction in risk was non-significant if they had not taken HRT within the past 5 years. The protective effect was dose related and similar results were noted whether the HRT was administered orally or via skin patches (but not by injection). For current users of HRT, the benefit of starting treatment 9 or more years after menopause was comparable to that of an earlier start.

Calcitonin reduces bone resorption by inhibiting the action of osteoclasts. It is now available as a nasal spray (Bergman, 1996). Increases in bone mass have been observed in women with osteoporosis who were treated with calcitonin (Consensus Development Conference, 1991; Riggs & Melton, 1992; Bergman, 1996). Its effects on cortical bone have not been established (Consensus Development Conference, 1991; Riggs & Melton, 1992). Calcitonin has an analgesic effect that may benefit patients with pain attributable to

vertebral fractures (Consensus Development Conference, 1991; Bergman, 1996). Salmon calcitonin is more potent than human calcitonin but patients may develop resistance in a relatively short period of time (Riggs & Melton, 1992). In women ages 68 to 72, Overgaard, Hansen, Jensen, and Christiansen (1992) observed dose-dependent increases in spinal bone mass over two years of treatment with nasal salmon calcitonin.

Bisphosphonates (including etidronate and FDA approved alendronate) impair the bone resorption activity of the osteoclasts. These compounds are retained within the bone for many years although it is unknown whether there are long-term increases in BMD associated with short-term treatment (Riggs & Melton, 1992; Bergman, 1996). They are poorly absorbed and may produce severe gastrointestinal irritation if the instructions for use are not followed exactly. Dosage is a critical factor because at high doses, the mineralization of newly synthesized bone matrix is impaired (Riggs & Melton, 1992).

Several studies have looked at the effects of alendronate therapy on BMD and Fracture risk. Liberman et al. (1995) pooled the results of 2 multicenter studies. The women in the study were all postmenopausal, between 45 and 80 years old, and with a BMD of the lumbar spine at least 2.5 standard deviations below the mean value for premenopausal white women. Forty percent of the subjects were assigned to receive a placebo, while 20 percent were assigned to each of the three treatment levels (5, 10, or 20 mg of alendronate/day). The treatment was administered for 2 years. During the third year, the women receiving 20 mg/day were switched to 5 mg/day based on a study that found 20 mg/day was more than was necessary. All women also received 500 mg/day of elemental calcium. DXA was used to evaluate BMD in the lumbar spine, femoral neck, trochanter, forearm, and total body. Lateral spine films, vertebral-height ratios, vertebral deformities, height, and fractures were also recorded. Significant increases in the BMD of all sites except the forearm were observed for all 3 of the alendronate groups while significant decreases were found in the placebo group. The number of new vertebral fractures was greater in the placebo group (22 of 355 women or 6.2 percent vs. 17 of 526 women or 3.2 percent). For non-vertebral fractures there was a trend toward a reduced number of fractures in the alendronate group although the number of ankle, foot, or toe; face or skull; shoulder; and clavicle or sternum fractures was higher. Overall, the alendronate was well tolerated.

Karpf et al. (1997) completed a meta-analysis of five randomized, placebo-controlled trials with at least 2 years of follow-up (including the study by Liberman et al. cited above). All of the studies enrolled women ages 42 to 85 years who were postmenopausal for at least 4 years. The women had lumbar spine BMD values (DXA) of at least 2 standard deviations below the mean for young adult women (with at least 2.5 standard deviations below the mean required in 3 studies). Exclusion criteria were similar across the 5 studies (although patients with prevalent fractures were excluded in 1 study) as was the protocol for assessing nonvertebral fractures. Dosages of alendronate ranged from 2.5 mg/day to 20 mg/day. All women were given 500 mg/day of calcium (with vitamin D, if deficient). An intention-to-treat analysis was done. Baseline data from the individual studies revealed no clinically meaningful differences between groups (placebo vs. treatment) within a study. Mean age ranged from 59 years to 71 years across the different studies. In each of the individual studies, significant (p-values were not reported in the meta-analysis) mean percent increases (from baseline) in BMD of the lumbar spine, femoral neck, trochanter, and total body were observed in the alendronate group (as compared to the placebo group). Sixty patients from the placebo group (total n=590) reported fractures (an overall rate of 4.45 patients with fractures per 100 patient-years). Seventy-three patients from the alendronate group (total n=1012) reported fractures (an overall rate of 3.26 patients with fractures per 100 patient-years). The estimated cumulative incidence of nonvertebral fracture after 3 years was 12.6% in the placebo group and 9.0% in the alendronate group (RR=0.71, 95% CI: 0.502-0.997, p=0.048). Relative risk values (for fracture in the

alendronate group compared to the placebo group) for the individual studies ranged from 0.34 to 0.91 (all of which were non-significant). In these analyses, a patient with more than one fracture was only counted once.

In the study presented by Cummings et al. (1998), 4,432 women with low femoral neck BMD ($0.68 g/cm^2$ or less) but no vertebral fracture were randomly assigned to receive either alendronate (5 mg/day for 2 years and 10 mg/day for the remainder of the study) or placebo. The BMD value of $0.68 g/cm^2$ was believed to correspond to at least 2 standard deviations below the mean for normal young adult women but with subsequent data from NHANES III it was discovered that the BMD value actually represents a level of about 1.6 standard deviations below the mean. Those with calcium intakes of 1000 mg/day or less were given supplementary calcium and vitamin D. The patients ranged in age from 55 to 80 years and had been post-menopausal for at least 2 years. The study excluded women with a variety of medical conditions (including recent ulcer, dyspepsia requiring daily treatment, significant renal or hepatic dysfunction, blood pressure above 210 mmHg systolic or above 105 mmHg diastolic, recent MI, or thyroid dysfunction). They also excluded women treated within the previous 6 months with either estrogen or calcitonin and women treated anytime with bisphosphonates or sodium fluoride. Estrogen was taken by 11% of the placebo group and 9% of the alendronate group at some time during the trial. At baseline, the groups did not differ with respect to age, BMD, history of fracture, self-rated health status, height, body mass index, calcium intake, or smoking status. A total of 4272 (96%) women were evaluated at the final follow-up (an average of 4.2 years after randomization). At the final follow-up, 83% of the placebo group and 81% of the alendronate group were still taking the study medication. Treatment with alendronate increased average BMD at the femoral neck, total hip, and lumbar spine when compared to the placebo group (all $p<0.001$). The percent increase in BMD was comparable regardless of initial BMD level. Clinical fractures, fractures diagnosed by a physician, were experienced by 14.1% of the placebo group and 12.3% of the alendronate group ($p=0.07$). The difference between treatment groups was significant only for fractures other than the hip, wrist, or spine ($p=0.02$). There was a significant interaction between the effect of treatment on the risk of clinical fracture and the initial BMD level ($p=0.01$). Women with a lower initial femoral neck BMD (2.5 standard deviations below the mean using the NHANES III data for reference) who were treated with alendronate experienced a significantly reduced risk of clinical fractures (compared to women treated with placebo). The relative hazard was 0.64 (95% CI: 0.50-0.82). Fracture risk was not significantly affected in women with BMD values greater than –2.5 standard deviations below the mean (relative hazard=1.08, 95%CI 0.87-1.35). Based on radiographs obtained at the final follow-up, alendronate reduced the overall risk for a new radiographic vertebral fracture by 44% ($p=0.001$). Final radiographs were obtained from 4134 participants (95% of those alive at the time of final follow-up). The percentages of patients with adverse experiences were similar in the two groups. In the placebo group, 10.2% experienced an adverse event leading to discontinuation of treatment; in the alendronate group the value was 9.9%. An upper gastrointestinal tract event was experienced by 47.2% of the placebo group and 47.5% of the alendronate group.

There is limited data available regarding the effect of combining HRT and bisphosphonates. Early post-menopausal women (Wimalawansa, 1995) and postmenopausal women with established osteoporosis (Wimalawansa, 1997) were randomly assigned to one of 4 groups: HRT, intermittent cyclical etidronate (ICE), HRT + ICE, or control. All of the patients received supplemental calcium. The women with established osteoporosis also received supplemental vitamin D. It should be noted that ICE is not FDA approved for the treatment of postmenopausal osteoporosis. In both studies, the follow-up was intended to be 4 years, with monitoring at 3 month intervals and testing of BMD at the lumbar spine and femoral neck at baseline, 2 years, and 4 years. In the study of women with established osteoporosis, telephone contact every 6 weeks and

spinal x-rays (to identify vertebral fractures) were added to the protocol. To be eligible for the study, the early post-menopausal women were to have had no prior spine, hip, or wrist fractures and were not to be taking any medications that would affect calcium metabolism. They had no prior treatment with HRT or bisphosphonates. The women with established osteoporosis (as determined by a vertebral fracture or a BMD of the spine that was 2 standard deviations or more below the mean for healthy 35 year old women) received no prior (since menopause) HRT, anabolic steroids, glucocorticoids, calcitonin, fluoride, or bisphosphonates.

In the study of early post-menopausal women, there were 58 women enrolled initially. Of the 58, 50 were evaluated at 2 years (86%) and 43 (74%) at 4 years. Withdrawals from the study were due to relocation (n=6), inability to tolerate the medication (n=6), and other medical problems (n=3). For lumbar spine BMD, the percent changes from baseline (at 4 years) were 6.8% for the HRT group, 6.8% for the ICE group, 10.9% for the HRT+ICE group, and –3.8% for the control group (all $p<0.01$). For femoral neck BMD, the percent changes from baseline (at 4 years) were 4.0% for the HRT group, 1.2% for the ICE group, 7.3% for the HRT+ICE group, and –5.0% for the control group (all $p<0.01$ except ICE group $p<0.05$). The increases in BMD were greater than in the group receiving the combined treatment than in the groups receiving either HRT or ICE alone ($p<0.05$ for spine BMD and $p<0.01$ for femoral neck BMD).

Similar results were observed for the women with established osteoporosis. A total of 72 women were randomized and 58 completed the study. Five patients withdrew because of estrogen-related adverse effects, 2 withdrew because of inability to tolerate the medication, and 5 withdrew because of other medical problems. There was one death and one loss to follow-up. For lumbar spine BMD, the percent changes from baseline (at 4 years) were 7.0% for the HRT group, 7.3% for the ICE group, 10.4% for the HRT+ICE group, and –2.5% for the control group (all $p<0.001$ except control $p<0.05$). For femoral neck BMD, the percent changes from baseline (at 4 years) were 4.8% for the HRT group, 0.9% for the ICE group, 7.0% for the HRT+ICE group, and –4.4% for the control group (all $p<0.01$ except ICE group $p<0.05$). For the lumbar spine BMD, the combined group experienced significantly greater changes than either the HRT or ICE groups at both 2 years ($p<0.01$) and 4 years ($p<0.05$). For the femoral neck BMD, the differences were only significant at 4 years ($p<0.05$ vs. HRT and $p<0.01$ vs. ICE). The rate of new vertebral fractures was 17 per 1,000 patient-years in the combined treatment group, 33 per 1,000 patient-years in the HRT group, 54 per 1,000 patient years in the ICE group, and 89 per 1,000 patient-years in the control group. The difference was not significant ($p=0.07$). The corresponding numbers of new fractures were 1, 2, 3, and 5.

In the study presented by Lindsay et al. (1999) 428 postmenopausal women who were diagnosed with osteoporosis (defined as a BMD of the lumbar spine or femoral neck at least 2 standard deviations below the mean and a BMD of the other site at least 1.5 standard deviations below the mean) and who were receiving HRT for at least one year before study entry were randomly assigned to receive either alendronate (10 mg/day) or placebo in addition to their HRT regimen. The mean duration of HRT use was approximately 10 years. Supplemental calcium and vitamin D were also provided. A total of 11 patients in the Alendronate + HRT group discontinued the study (5 due to adverse events, 4 at the request of the patient, and 2 lost to follow-up). Twenty-three patients in the HRT only group discontinued the study (11 due to adverse events, 6 at the request of the patient, and 6 for other reasons). The two groups were similar at baseline with the exception of greater family history of osteoporosis and greater history of smoking among patients in the alendronate + HRT group. Repeat BMD testing of the lumbar spine at 6 and 12 months showed significant differences ($p<0.001$) between groups in the percent change from baseline. At 12 months, the alendronate + HRT group experienced a 3.6% increase while the HRT only group experienced a 1.0% increase. Similar values were observed for the hip trochanter (2.7% vs. 0.5%; $p<0.001$). The groups did not differ at the

femoral neck site (1.7% vs. 0.8%; p=0.07). The overall incidence of clinical adverse effects was similar in each group. The number of patients experiencing fractures during the study period was greater in the alendronate + HRT group but the difference was not significant (15 patients vs. 9 patients; p=0.29). There were no hip fractures or symptomatic vertebral fractures during the study period.

Raloxifene has recently been approved by the FDA for the prevention of osteoporosis. Raloxifene is a selective estrogen receptor modulator that functions to increase BMD (Delmas et al., 1997; Lufkin et al., 1998; Khovidhunkit & Shoback, 1999). However, the changes in density are not as great as would be expected with either HRT (Lufkin et al., 1998) or bisphosphonates. As a result of the lesser effect on BMD, a longer time interval between testing would be required to accurately reflect change.

Ettinger et al. (1999) reported data from 7,705 women, all of whom were at least 2 years postmenopausal and diagnosed with osteoporosis based on either low BMD (WHO criteria) or radiographically apparent fractures. Among other criteria, the study excluded those who had taken an androgen, calcitonin, or bisphosphonate within the previous 6 months; oral estrogen within the previous 2 months; fluoride therapy for more than 3 months during the previous 2 years; or those who had undergone systemic glucocorticoid therapy for more than 1 month within the past year. Eligible patients were randomized to receive either raloxifene (60 mg or 120 mg) or placebo. All of the women took calcium and vitamin D supplements in addition to the study medication. Vertebral radiographs were taken at baseline, 24 months, and 36 months (with more frequent examinations if there were symptoms of fracture). Non-vertebral fractures were determined by questioning the patients at 6 month intervals (traumatic fractures and fractures of the fingers, toes, and skull were excluded). Spine and femoral neck BMD were measured annually. Patients were required to discontinue participation in the study if they had an excess loss of BMD or more than 2 incident vertebral fractures. Based on the data from 6,828 patients with baseline and follow-up radiographs, there was a lower incidence of new vertebral fractures in the raloxifene groups (6.6% of the 60 mg/day group and 5.4% of the 120 mg/day group vs. 10.1% of the placebo group) with no difference between the two dosage groups except in the subgroup of patients with prior fractures. The total relative risk was 0.7 (95%CI: 0.6-0.9) for the 60 mg/day group and 0.6 (95%CI: 0.4-0.7) for the 120 mg/day group. The incidence of non-vertebral fractures was similar (9.3% in the placebo group and 8.5% in the overall raloxifene group). BMD at 36 months was increased by 2.1% at the femoral neck and 2.6% at the spine in the 60 mg group and 2.4% and 2.7%, respectively, in the 120 mg group (all $p<0.001$ compared to the placebo group). The percentage of patients in each group that withdrew from the study for multiple fractures or for excessive BMD loss was 3.6% for the placebo group, 1.1% for the 60 mg group, and 0.9% for the 120 mg group. The only serious adverse effects believed to be causally related to raloxifene were venous thromboembolic events (reported in 0.3% of the placebo group, 1.0% of the 60 mg group and 1.0% of the 120 mg group). Breast cancer was less frequent in the raloxifene groups (RR=0.3; 95%CI: 0.2-0.6) while endometrial cancer was comparable (4 patients in the placebo group, 4 in the 60 mg group, and 2 in the 120 mg group).

Therefore, although there is evidence of a reduced risk of vertebral fracture with raloxifene, there is no evidence to support a reduced risk of hip or wrist fractures. The Multiple Outcomes of Raloxifene Evaluation (MORE) trial as described by Ettinger et al. (1999), above, was not specifically powered to detect a difference in hip fractures.

Adequate calcium intake at all stages of life is necessary to assure bone growth and to achieve peak bone mass (Consensus Development Conference, 1991). There is some evidence that calcium supplements (up to 2000 mg daily) can reduce the rate of bone loss after the menopause (Riggs & Melton, 1992; Alexeeva et al., 1994). Women who have a

low base-line calcium intake, older women, and women with osteoporosis may receive the greatest benefit from supplemental calcium (Riggs & Melton, 1992). Several studies have established a relationship between calcium intake and a reduced risk of fracture (Alexeeva et al., 1994).

Stimulators of bone formation include fluoride, parathyroid hormone, growth factors, and other modalities such as vitamin D and exercise. Fluoride stimulates osteoblasts and increases cancellous bone mass (Consensus Development Conference, 1991). The increased bone mass following fluoride therapy is abnormal in structure and may result in decreased bone strength (Kaltenborn, 1992; Riggs & Melton, 1992) and cortical bone might be compromised (Riggs & Melton, 1992). Gastric irritation and peripheral bone pain are common side effects (Riggs & Melton, 1992; Consensus Development Conference, 1993). Riggs, Melton, and O'Fallon (1996) reanalyzed data from clinical trials with fluoride and estrogen therapy and found that high bone turnover was an independent risk factor for fracture. Estrogen therapy reduced bone turnover and prevented vertebral fractures but some patients on high doses of fluoride experienced an increase in bone turnover and an increased risk of fracture. Reginster et al. (1998) randomly assigned 200 women with lumbar spine BMD values of 2.5 standard deviations or more below the mean to receive either fluoride and calcium or calcium alone. They excluded patients with hip fracture but not those with vertebral fractures although only 7 patients had vertebral fractures at the time of enrollment. Based on data from 164 patients with follow-up radiographs at from 1 to 4 years after randomization, there were fewer new vertebral fractures in the fluoride plus calcium group than in the calcium only group (2.4% vs. 10%; p=0.05). Thus, while fluoride has been shown to be a promising therapy in some trials, it remains under investigation.

Parathyroid hormone, when given intermittently in low doses, appears to stimulate bone formation. This is in contrast to high plasma concentrations or continuous infusion which promote bone resorption (Riggs & Melton, 1992; Consensus Development Conference, 1993). Cortical bone may be lost while cancellous bone is gained (Riggs & Melton, 1992).

Growth factors, if they can be properly targeted to affect only bone, have the potential to stimulate bone formation. An alternative would be to use drugs that stimulate osteoblasts to produce growth factors (Riggs & Melton, 1992). However, these approaches are still investigational.

Vitamin D deficiency is common in the elderly. With insufficient amounts of vitamin D, increased bone turnover and cortical bone loss are possible (as a result of secondary hyperparathyroidism) and newly formed bone may be less well mineralized (Alexeeva et al., 1994).

Inactivity has been identified as a risk factor for osteoporosis but the exact nature of an exercise program to preserve bone health has not been established (Alexeeva et al., 1994; Bergman, 1996). However, exercise in the elderly has been shown to prevent bone loss and promote mobility, agility, and muscle strength (Consensus Development Conference, 1993).

Many fractures among the elderly are a result of a fall. Minimizing the risk (or consequences) of falling may be effective in reducing the number of fractures, although the results of clinical trials have been inconsistent in this regard. Aging, medications, and disease all contribute to reduced balance, muscle strength, and agility (Consensus Development Conference, 1991). Home environmental hazards should be eliminated. Wearing special garments (hip protectors) to cushion falls was proposed as a way to reduce the number of hip fractures (Consensus Development Conference, 1993).

Although the use of BMD measurements to identify individuals at risk for osteoporotic fracture is well established, there are a number of issues that cannot be answered at this

point in time. Some of these include where to place the intervention threshold, duration of various therapeutic interventions, use of combination drug regimens, and the ultimate goal of therapy (reduction of all fractures or just hip fractures).

The Program

A screening program requires facilities and trained personnel for diagnosis and treatment and must be cost-effective (Alexeeva et al., 1994). Repa-Eschen (1993) stated that by identifying osteoporosis at an earlier, more treatable stage, the costs associated with treatment and morbidity would be reduced.

Risks and Limitations

Bone densitometry is a non-invasive procedure and is considered to be safe and generally acceptable (Cummings & Black, 1986). SPA, DPA, DXA, QCT and ultrasound densitometry have all been approved by the FDA. The radiographic equipment for RA is regulated by the FDA (Erlichman & Holohan, 1995). Radiation doses for SPA, SXA, and DXA are low and, while DPA doses are somewhat higher, all values are well below those of other diagnostic x-rays. A single-energy QCT scan produces the highest dose of the commonly used methods. The QCT dose equivalent is comparable to a dental x-ray (Genant et al., 1993).

Rubin and Cummings (1992) surveyed 260 women who had experienced densitometry of the spine (58% also had the hip tested). Ninety-two percent thought the test was worthwhile, 70 percent indicated that they would like to be tested again in the future, and 90 percent reported they would encourage other women their age to have the test.

As described previously in this report, each of the common measurement techniques has limitations. SPA and SXA are limited to peripheral sites and cannot differentiate cortical and trabecular bone. DPA and DXA can be used for spine measurements but also cannot distinguish cortical and trabecular bone. QCT can be used at all measurement sites (although most frequently for the spine) and can estimate separate cortical and trabecular densities but the radiation doses are higher. RA has been developed primarily for use with the hand. Ultrasound has not been validated and its precision and accuracy are uncertain. Accuracy and precision error values must be considered when determining the appropriate measurement system for a particular application.

Alternative Forms of Treatment

SPA, SXA, DPA, DXA, QCT, Ultrasound and RA as described previously in this report, are all physical measurements of skeletal mass (Alexeeva et al., 1994). Additional physical measurements include bone biopsy, conventional radiography (including the Singh Index), radiogrammetry, Compton (photon) scattering, neutron activation, and biochemical analyses for markers of bone resorption and formation. A brief explanation of each follows.

A bone biopsy is an invasive procedure that allows for bone mass to be evaluated while osteomalacia and certain forms of secondary osteoporosis can be excluded (Office of Medical Applications, 1984; Kaltenborn, 1992). By administering a timed cycle of tetracycline before the biopsy, it is possible to determine the rate of bone turnover (Kaltenborn, 1992). The procedure is considered safe but does require specialized equipment and expert analysis (Office of Medical Applications, 1984; Davis, 1987). With the development of precise noninvasive techniques, biopsy is now a secondary procedure in establishing the diagnosis of osteoporosis (Kaltenborn, 1992).

Conventional skeletal radiography may reveal defects in the upper and lower surfaces of involved vertebrae, one or more vertebral crush fractures, and diffuse demineralization and cortical thinning of long bones (Davis, 1986; Davis, 1987). It is an insensitive indicator of osteoporosis; a high degree of bone loss (20 to 30 percent or more) will have occurred before osteoporosis is indicated (Office of Medical Applications, 1984; Kaltenborn, 1992). The Singh Index (SI), a method of grading the trabecular bone in the femoral neck on a 1 to 6 scale, has been useful in epidemiological studies of hip fracture but less valuable in young, healthy women (Alexeeva et al., 1994). Masud, Jawed, Doyle, and Spector (1995) compared SI scores with DXA bone density values and found that the SI was simple and reproducible with low sensitivity but relatively high specificity. They recommended that anyone with a low SI grade (4 or less) should be considered for bone densitometry.

The cortical thickness of metacarpal and other tubular bones can be determined by radiogrammetry. Using standard radiographs, the width of the marrow cavity is subtracted from the outer bone diameter. Although the technique is limited to measurements of the appendicular skeleton, spine and hip fractures are uncommon among women with normal cortical thickness (Davis, 1986).

Trabecular bone density at peripheral sites can be measured by a method known as Compton (or photon) scattering (Davis, 1987; Alexeeva et al., 1994). The electron density (total density of bone plus marrow) of the scattering medium determines the extent of scattered radiation. Density can be determined with an accuracy and precision of about 3 percent but the method hasn't been adequately evaluated as a screening tool.

Neutron activation is a research technique that, until recently, was the only way to determine total body calcium content (Alexeeva et al., 1994). Local calcium content can be measured, as well. The body is irradiated with neutrons. Many elements become radioactive, including calcium. The decay back to a non-radioactive state is measured with a counter (Davis, 1986). The procedure is based on the assumption that 99 percent of the calcium in the body is in the skeleton so extraosseous calcification detracts from the method's validity (Davis, 1987, Alexeeva et al., 1994).

Magnetic resonance (MR) imaging has been proposed as a way to quantitate bone microarchitecture as well as bone density (Lang, Augat, Majumdar, Ouyang, & Genant, 1998; Grampp, Henk, & Imhof, 1999). Link et al. (1998) found that T2* decay at the proximal femur was fastest in premenopausal subjects and slowest in postmenopausal patients with fractures (all sites $p<0.01$ except non-significant difference at trochanteric region). BMD and T2* measurements at the femoral neck, Ward's triangle, intertrochanteric region, and total femur were significantly correlated (all $r\geq0.43$; $p<0.05$). T2* of the Ward's triangle region was the strongest discriminator between postmenopausal control subjects (age-matched women with no fractures and no history suggestive of metabolic bone disease) and postmenopausal patients with fractures (OR=3.02, 95%CI: 1.6-5.8). Kang, Paley, Ordidge, & Speller (1999) observed differences in calcaneal T2*, BMD, and QUS measures between postmenopausal women with low BMD and those with normal BMD. The correlation between the T2* and BMD measures was significant ($r=-0.80$, $p<0.0001$). The *in vivo* precision of MR imaging was poorer than that of DXA or QUS.

Biochemical markers of bone resorption and formation represent another potential clinical tool (ICSI, 2000). They cannot be used to diagnose osteopenia but may be useful in monitoring the effects of therapeutic intervention in patients with osteoporosis (Riis, 1993; Whitaker, 1996). It has also been suggested that biochemical markers may also be useful in diagnosing different types of osteoporosis (such as high turnover vs. low turnover) (Kleerekoper & Edelson, 1996). The most efficient markers of bone formation are osteocalcin (also known as bone Gla-protein or BGP) and bone alkaline phosphatase while the most efficient markers of bone resorption are pyridinoline and deoxypyridinoline collagen crosslinks and urinary aminoterminal telopeptide (NTX) of type I collagen (Delmas, 1993; Russell, 1997). Many other markers are being investigated (Whitaker,

1996; Christenson, 1997). Combining the rate of bone loss with bone mass measurements might enhance the prediction of fracture risk (Riis, 1993) although Alexeeva et al. (1994) concluded that there was no evidence that the rate of loss contributes to fracture risk independently of low bone mass. The sensitivity and specificity of each of the markers of bone loss is variable but low (Alexeeva, 1994).

The use of risk factors to predict low bone mass and fracture risk was evaluated by Ribot et al. (1995). Both univariate models and multivariate models were considered. The models developed to predict bone mass have coefficient of determination (r^2) values of less than 0.43 indicating that much of the variance in bone mass is not explained by the risk factors. Models to predict fracture risk were poorer. The authors concluded the poor performance of the models was due to unmeasured factors, especially genetic effects.

Cummings et al. (1995) evaluated a wide range of potential risk factors in a sample of 9516 women age 65 and above. Data on atraumatic hip fractures were collected during an average 4.1 year follow-up. Some factors that were significant predictors of hip fracture in univariate models were not significant in multivariate models. The presence of 5 or more risk factors was associated with a higher incidence of hip fracture. It was suggested that the assessment of risk factors and the measurement of bone density have a complementary value in the prediction of hip fractures. A small number of women with multiple risk factors and low bone density have an especially high fracture risk.

Epidemiology and Costs

Osteoporosis affects 25 to 35 million Americans. In the United States in 1995, there were 432,448 hospitalizations with a primary diagnosis of osteoporotic fracture, 179,221 nursing home stays attributed to osteoporotic fractures, and 3.4 million outpatient physician, outpatient hospital, and emergency room examinations related to osteoporotic fractures among people age 45 and older. The annual cost of treatment for osteoporotic fractures in this population was reported to be $13.8 billion (Ray, Chan, Thamer, & Melton, 1997). Mortality rates range from 17 to 24 percent during the first 12 months following admission to a hospital for a hip fracture. This represents a 5 to 20 percent reduction from expected survival (Repa-Eschen, 1993). For hip fracture patients there is a 25 percent chance of long-term institutionalization and a less than 50 percent chance of full recovery (Resnick & Greenspan, 1989). In Rochester, Minnesota, the 30-day case fatality rate following first hip fracture (based on data from 1262 patients between 1953 and 1982) was 7.2% overall (men and women). For women, the rate was 5.4% while for men the rate was 14.4%. The values were higher for patients age 75 and over as compared to those under 75 years of age (Poor, Jacobsen & Melton, 1994).

In Minnesota, it is estimated that 32% of non-Hispanic, white women over age 50 have osteoporosis as defined by the WHO criteria of more than 2.5 standard deviations below the mean for young, normal women (Table 4). For the hip alone, the rate is 24%. For other populations, the estimated prevalences of osteoporosis of the hip are as follows: black women, 9%; Asian/Indian women, 17%; and Hispanic women, 12%. The estimated total prevalence of hip osteoporosis among women over age 50 in Minnesota is 23%.

Table 4. Estimated Number of Affected Minnesotans (prevalence)

	Non-Hispanic White		Black		Asian/ Indian		Hispanic		Total	
	N	% *	N	% *	N	% *	N	% *	N	% *
Osteoporosis Prevalence										
Any site										
Men										
Women	199,012	32[a]								
Hip alone										
Men	33,618	6[b]	140	2[c]	411	5[d]	119	3[e]	34,288	6
Women	148,419	24[f]	555	9[g]	1,750	17[h]	485	12[i]	151,209	23
Annual Fracture Incidence										
Hip										
Men	1,318	0.3[j]	13	0.2[k]	5	0.1[l]	3	0.1[m]	1,339	0.3
Women	3,699	0.6[j]	9	0.1[k]	19	0.2[l]	8	0.2[m]	3,735	0.6
Spine										
Men	1,073	0.2[n]								
Women	3,607	0.6[n]								
Forearm										
Men	858	0.2[o]								
Women	4483	0.7[o]								

*Percentage of those ≥ 50 years of age

[a]Prevalence rates from Rochester (Melton, 1995)

[b]Prevalence rates from Hologic normal data from men (Alexeeva et al., 1994)

[c]Prevalence rates assumed to be one-half those for white men

[d]Prevalence rates assumed to be the same as those for white men

[e]Prevalence rates assumed to be two-thirds of those for white men

[f]Prevalence rates from NHANES III (Looker, Johnston, et al., 1995a)

[g]Prevalence rates assumed to be one-half of those for white women (overall prevalence is one-half)

[h]Prevalence rates assumed to be the same as those for white women (no NHANES data reported)

[i]Prevalence rates assumed to be two-thirds of those for white women (overall prevalence is two-thirds)

[j]Hip fracture incidence rates from Rochester (Madhok, Melton, Atkinson, O'Fallon, & Lewallen, 1993)

[k]Hip fracture incidence rates from Washington, D.C. (Farmer, White, Brody, Bailey, 1984)

[l]Hip fracture incidence rates from Okinawa (Ross, Noimatsu, et al., 1991)

[m]Hip fracture incidence rates from Texas (Bauer, 1988)

[n]Vertebral fracture incidence rates from Rochester (Cooper, Atkinson, O'Fallon, & Melton, 1992)

[o]Forearm fracture incidence rates from Rochester (Owen, Melton, Johnson, Ilstrup, & Riggs, 1982)

Chrischilles, Shireman, and Wallace (1994) modeled the lifetime fracture impact by age cohort with 10,000 women per cohort. For 10,000 women with an initial age of 50, it was estimated that 3911 will experience 5419 fractures. The total cost was estimated at $24.1 million. Projecting to the U.S. population of white postmenopausal women, it was estimated that over a 10-year period there would be 5.2 million fractures, 2 million person-years of fracture-related functional impairment, and $45.2 billion total direct costs of fracture-related care.

Annual costs for the various treatment options (derived from the Redbook Average Wholesale Prices) are as follows.

Calcium and Vitamin D	$20
Hormone Replacement Therapy	$180
Raloxifene	$740
Alendronate	$695
Calcitonin (Nasal)	$735

Summary

With respect to densitometry as a diagnostic tool for the identification and treatment of osteoporosis, the ICSI Technology Assessment Committee finds the following:

1. Osteoporosis is a significant health issue. Osteoporosis-related fractures result in reduced quality of life and significant annual costs for treatment.

2. The value of preventive interventions such as adequate calcium intake, vitamin D supplementation, and weight bearing physical activity has been proven. There is a range of possible interventions some of which can be initiated without screening for bone density. The available options should be thoroughly explained to all patients. However, many patients exist and will continue to exist who present with osteoporosis; identification and treatment of those patients is the focus of this report.

3. Single-photon absorptiometry (SPA), single-energy x-ray absorptiometry (SXA), dual-photon absorptiometry (DPA), dual-energy x-ray absorptiometry (DXA) (central and peripheral; conventional and high-speed systems), quantitative computed tomography (QCT) (central and peripheral), and quantitative ultrasound (QUS) are safe procedures.

4. DXA is most commonly used because it allows measurement of sites (e.g., spine and hip) that are of greatest interest. DXA is also preferred because of its ability to measure multiple sites, its relatively low radiation exposure, and its superior precision. As a result of the high precision, DXA is the only technique suitable for the repeat testing necessary to assess the effects of treatment. DXA is more expensive and less accessible than QUS systems. Recent advances in DXA technology including peripheral DXA and fan beam DXA have facilitated measurement of peripheral sites and reduced scan times, respectively. However, neither of these new techniques offers as high a precision as has been found with conventional DXA systems. Quantitative ultrasound (QUS) offers greater administrative flexibility and lower cost than DXA but, at present, there is no agreement on how to interpret the data for the purpose of diagnosing osteoporosis and the measurement is less precise than DXA. Prospective studies have shown that QUS can be used to predict hip fracture risk in women 65 years of age and older. QUS may serve as an inexpensive "quasi-screening" test to identify those for whom more precise measurement (with DXA) would be of value. QUS may also be used to encourage women toward therapy. However, the precision of QUS does not allow for repeat measurements within the 2 to 3 year window appropriate with current DXA systems. It is important to note that

T-scores from different modes of testing (i.e., DXA, QUS) cannot be compared directly. (Conclusion Grade II, See Appendix)

5. At present, data from the different measurement techniques is based on different normative data bases. There is a need for a single normative data base against which all of the techniques can be evaluated.

6. *At present, there is insufficient evidence to support mass screening for bone mineral density (BMD). However, BMD testing can provide important information for selected patients and therefore the need for bone mineral density (BMD) testing must be determined on an individual patient basis.* A comprehensive analysis from the National Osteoporosis Foundation has resulted in nomograms to use in determining an individual's need for diagnostic testing and/or treatment. The nomograms take into account the patient's age, risk factors, prior fracture, and prior treatment. The decision whether to test for BMD should be made in the context of the treatment being considered and only after review of the patient's current health status. At menopause, all women should be counseled about the use of HRT. Women who are not able to take estrogen should be counseled on alternative treatments. *Testing is of value when making individual decisions about therapies in lieu of HRT as well as when an individual's decision about HRT would be influenced by her knowledge of her BMD. Testing for bone density is contraindicated if the results of the testing would not influence the patient's treatment decision.* The recent NOF recommendations call for BMD testing for women who have been on HRT for prolonged periods. The value of this testing is in detecting treatment failures. Recent information suggests that, for patients with low bone mass despite long-term HRT, the addition of bisphosphonates can increase bone mass. It would seem logical for there to be similar concern over possible treatment failures with other therapies.

7. The value of repetitive testing for either monitoring the effects of therapy or for making decisions about interventions in elderly women is poorly understood. Given the precision errors of the current technologies, yearly densitometry is not clinically indicated. Using DXA, follow-up intervals of at least two years are needed to reliably detect changes in bone mass. A two year interval will also avoid scans that suggest non-response to treatment in women who are free of significant risk factors such as glucocorticoid use or metabolic disease. However, patients in higher risk categories (such as those requiring glucocorticoid therapy, those with very low bone mass, those with fracture despite what is assumed to be effective therapy, or those with osteopenia or osteoporosis who are early after solid organ or allogeneic bone marrow transplantation) may need more frequent scans. Decisions about when to repeat testing must be individualized based on the patient's overall health status, risk status (including corticosteroid dosage), and baseline BMD value.

8. Each clinical center should establish their own short-term precision values. Although there is published data on CV values for the different techniques, the actual values observed at a given site may vary.

References

Evidence is classed and graded as described below.

I. CLASSES OF RESEARCH REPORTS

Primary Reports of New Data Collection:

Class A:	Randomized, controlled trial
Class B:	Cohort study
Class C:	Non-randomized trial with concurrent or historical controls Case-control study

Study of sensitivity and specificity of a diagnostic test
Population-based descriptive study

Class D: Cross-sectional study
Case series
Case report

Reports that Synthesize or Reflect upon Collections of Primary Reports:

Class M: Meta-analysis
Decision analysis
Cost-benefit analysis
Cost-effectiveness study

Class R: Review article
Consensus statement
Consensus report

Class X: Medical opinion

II. CONCLUSION GRADES

Key conclusions (as determined by the work group) are supported by a conclusion grading worksheet that summarizes the important studies pertaining to the conclusion. Individual studies are classed according to the system defined in Section I, above, and are assigned a designator of +, -, or ø to reflect the study quality. Conclusion grades are determined by the work group based on the following definitions:

Grade I: The evidence consists of results from studies of strong design for answering the question addressed. The results are both clinically important and consistent with minor exceptions at most. The results are free of serious doubts about generalizability, bias, and flaws in research design. Studies with negative results have sufficiently large samples to have adequate statistical power.

Grade II: The evidence consists of results from studies of strong design for answering the question addressed, but there is uncertainty attached to the conclusion because of inconsistencies among the results from different studies or because of doubts about generalizability, bias, research design flaws, or adequacy of sample size. Alternatively, the evidence consists solely of results from weaker designs for the question addressed, but the results have been confirmed in separate studies and are consistent with minor exceptions at most.

Grade III: The evidence consists of results from a limited number of studies of weak design for answering the question addressed. Evidence from studies of strong design is either unavailable because no studies of strong design have been done or because the studies that have been done are inconclusive due to lack of generalizability, bias, design flaws, or inadequate sample sizes.

Grade IV: The support for the conclusion consists solely of the statements of informed medical commentators based on their clinical experience, unsubstantiated by the results of any research studies.

The symbols **+**, **–**, **ø**, and **N/A** found on the conclusion grading worksheets are used to designate the quality of the primary research reports:
+ indicates that the report has clearly addressed issues of inclusion/exclusion, bias, generalizability, and data collection and analysis;
– indicates that these issues have not been adequately addressed;
ø indicates that the report is neither exceptionally strong or exceptionally weak;
N/A indicates that the report is not a primary reference and therefore the quality has not been assessed.

AACE clinical practice guidelines for the prevention and treatment of postmenopausal osteoporosis. *Endocr Pract* 1996;2:157-171. (Class R)

Alexeeva L, Burkhardt P, Christiansen C, et al. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. *World Health Organization Technical Report Series 843*. Geneva: World Health Organization. 1994. (Class R)

Alhava EM. Bone density measurements. *Calcif Tissue Int* 1991;49(Suppl):S21-S23. (Class R)

Baran DT. Osteoporosis: Monitoring techniques and alternate therapies. *Obstet Gynecol Clin North Am* 1994;21:321-335. (Class R)

Bauer RL. Ethnic differences in hip fracture: a reduced incidence in Mexican Americans. *Am J Epidemiol* 1988;127:145-149. (Class C)

Bauer DC, Glüer CC, Cauley JA, et al. For the Study of Osteoporotic Fractures Research Group. Broadband ultrasound attenuation predicts fractures strongly and independently of densitometry in older women. *Arch Intern Med* 1997;157:629-634. (Class B)

Bergman D. Osteoporosis: an aggressive approach for prevention and management. *Endocr Pract* 1996;2:20-23. (Class R)

Black DM. Why elderly women should be screened and treated to prevent osteoporosis. *Am J Med* 1995;98(2A):67S-75S. (Class R)

Chrischilles E, Shireman T, Wallace R. Costs and health effects of osteoporotic fractures. *Bone* 1994;15:377-386. (Class M)

Christenson RH. Biochemical markers of bone metabolism: an overview. *Clin Biochem* 1997;30:573-593. (Class R)

Christiansen C. Postmenopausal bone loss and the risk of osteoporosis. *Osteoporosis Int* 1994;Suppl 1:547-551. (Class R)

Comptson JE, Cooper C, Kanis JA. Bone densitometry in clinical practice. *BMJ* 1995, 310:1507-1510. (Class R)

Consensus development conference: diagnosis, prophylaxis, and treatment of osteoporosis. *Am J Med* 1993;94:646-649. (Class R)

Consensus development conference: prophylaxis and treatment of osteoporosis. *Am J Med* 1991;90:107-110. (Class R)

Cooper C, Atkinson EJ, O'Fallon WM, Melton LJ III. Incidence of clinically diagnosed vertebral fractures: a population based study in Rochester, Minnesota 1985-1989. *J Bone Miner Res* 1992;7:221-227. (Class C)

Cosman F, Herrington B, Himmelstein S, Lindsay, R. Radiographic absorptiometry: a simple method for determination of bone mass. *Osteoporosis Int* 1991;2:34-38. (Class D)

Cummings SR, Black DM. Should perimenopausal women be screened for osteoporosis? *Ann Intern Med* 1986;104:817-823. (Class R)

Cummings SR, Black DM, Nevitt MC, et al. Appendicular bone density and age predict hip fracture in women. *JAMA* 1990;263:665-668. (Class B)

Cummings SR, Black DM, Nevitt MC, et al. Bone density at various sites for prediction of hip fractures. *Lancet* 1993;341:72-75. (Class B)

Cummings SR, Black DM, Thompson DE, et al. for the Fracture Intervention Trial Research Group. Effect of alendronate on risk of fracture in women with low bone density but without vertebral fractures. *JAMA* 1998;280:2077-2082. (Class A)

Cummings SR, Nevitt MC, Browner WS, et al. Risk factors for hip fracture in white women. *N Engl J Med* 1995;332:767-773. (Class B)

Davis MR. Osteoporosis detection-screening the screens. *Contemporary Ob/Gyn* 1986;95-106. (Class R)

Davis MR. Screening for postmenopausal osteoporosis. *Am J Obstet Gynecol* 1987;156:1-5. (Class R)

Davis JW, Vogel JM, Ross PD, Wasnich RD. Disease vs etiology: the distinction should not be lost in the analysis. *J Nucl Med* 1989;30:1273-1276. (Class X)

Delmas PD. Biochemical markers of bone turnover I: Theoretical considerations and clincal use in osteoporosis. *Am J Med* 1993;95(suppl 5A):11S-16S. (Class R)

Delmas PD, Bjarnason NH, Mitlak BH, et al. Effects of raloxifene on bone mineral density, serum cholesterol concentrations, and uterine endometrium in postmenopausal women. *N Engl J Med* 1997;337:1641-1647. (Class A)

Department of Health and Human Services. Medicare coverage of and payment for bone mass measurements. *Federal Register* 1998;63:34320-34328. (Class not assignable)

Düppe H, Gärdsell P, Nilsson B, Johnell O. A single bone density measurement can predict fractures over 25 years. *Calcif Tissue Int* 1997;60:171-174. (Class B)

Erlichman M, Holohan TV. Bone densitometry: patients with asymptomatic primary hyperparathyroidism. *Health Technology Assessment Report*, 1995, No. 6. Rockville, MD: U.S. Department of Health and Human Service, Agency for Health Care Policy and Research. (Class R)

Ettinger B, Black DM, Mitlak BH et al. for the Multiple Outcomes of Raloxifene Evaluation (MORE) Investigators. Reduction of vertebral fracture risk in postmenopausal women with osteoporosis treated with raloxifene. *JAMA* 1999;282:637-645. (Class A)

Farmer ME, White LR, Brody JA, Bailey KR. Race and sex differences in hip fracture incidence. *Am J Public Health* 1984;74:1374-1380. (Class C)

Faulkner KG. Bone densitometry: choosing the proper skeletal site to measure. *J Clin Densit* 1998;1:279-285. (Class R)

Franck H, Munz M, Scherrer M. Bone mineral density of opposing hips using dual energy X-Ray absorptiometry in single-beam and fan-beam design. *Calcif Tissue Int* 1997;61:445-7. (Class D)

Gärdsell P, Johnell O, Nilsson BE. The predictive value of bone loss for fragility fractures in women: A longitudinal study over 15 years. *Calcif Tissue Int* 1991;49:90-94. (Class B)

Garnero P, Shih WJ, Gineyts E, Karpf DB, Delmas PD. Comparison of new biochemical markers of bone turnover in late postmenopausal osteoporotic women in response to alendronate treatment. *J Clin Endocrinol Metab* 1994;79:1693-1700. (Class A)

Genant HK, Engelke K, Fuerst T, et al. Noninvasive assessment of bone mineral and structure: state of the art. *J Bone Miner Res* 1996;11:707-730. (Class R)

Genant HK, Faulkner KG, Glüer CC, Engelke K. Bone densitometry: current assessment. *Osteoporosis Int* 1993;Suppl 1:S91-97. (Class R)

Glüer CC for the International Quantitative Ultrasound Consensus Group. Quantitative ultrasound techniques for the assessment of osteoporosis: expert agreement on current status. *J Bone Miner Res* 1997;12:1280-1288. (Class R)

Glüer CC, Jergas M, Hans D. Peripheral measurement techniques for the assessment of osteoporosis. *Semin Nucl Med* 1997;27:2290247. (Class R)

Grampp S, Henk CB, Imhof H. CT and MR assessment of osteoporosis. *Semin Ultrasound CT MR* 1999;20:2-9. (Class R)

Greenspan SL, Maitland-Ramsey L, Myers E. Classification of osteoporosis in the elderly is dependent on site-specific analysis. *Calcif Tissue Int* 1996;58:409-414. (Class D)

Hans D, Dargent-Molina P, Schott AM, et al., for the EPIDOS prospective study group. Ultrasonographic heel measurements to predict hip fracture in elderly women: The EPIDOS prospective study. *Lancet* 1996;348:511-514. (Class B)

Hassager C, Jensen SB, Godfredsen A, Christiansen C. The impact of measurement errors in diagnostic value of bone mass measurement: theoretical considerations. *Osteoporosis Int* 1991;1:250-256. (Class R)

Hayes, Inc. Ultrasound versus dual energy x-ray absorptiometry for osteoporosis diagnosis and screening. 1998. (Class R)

Hoshi K, Yamada H, Tsukikawa S, Sugano T, Endo C, Sato A. The bone mineral density change with aging of Japanese women measured by single energy X-ray absorptiometry and dual energy X-ray absorptiometry. *Arch Gynecol Obstet* 1993;253:65-69. (Class D)

Huang C, Ross PD, Yates AJ, et al. Prediction of fracture risk by radiographic absorptiometry and quantitative ultrasound: a prospective study. *Calcif Tissue Int* 1998;63:380-384. (Class B)

Hui SL. Baseline measurement of bone mass predicts fracture in white women. *Ann Intern Med* 1989;111:355-361. (Class B)

Institute for Clinical Systems Integration. Technology Assessment #9: Screening Tests. Minneapolis: ICSI. 1993. (Class R)

Institute for Clinical Systems Improvement. Technology Assessment (in development): Biochemical Markers of Bone Turnover in Osteoporosis. Minneapolis: ICSI. 2000. (Class R)

Johnston CC, Melton LJ III. Bone densitometry. In: Riggs BL and Melton LJ III (eds.) *Osteoporosis: Etiology, Diagnosis, and Management, Second Edition*. Philadelphia: Lippincott-Raven. 1995:275-297. (Class R)

Johnston CC, Slemenda CW, Melton LJ III. Clinical use of bone densitometry. *N Engl J Med* 1991;324:1105-1109. (Class R)

Kaltenborn KC. Perspectives on osteoporosis. *Clin Obstet Gynecol* 1992;35:901-912. (Class R)

Kang C, Paley M, Ordidge R, Speller R. In vivo MRI measurements of bone quality in the calcaneus: a comparison with DXA and ultrasound. *Osteoporosis Int* 1999;9:65-74. (Class C)

Kanis JA, Melton LJ III, Christiansen C, Johnston CC, Khaltaev N. The diagnosis of osteoporosis. *J Bone Miner Res* 1994;8:1137-1141. (Class R)

Karpf DB, Shapiro DR, Seeman E, et al. for the Alendronate Osteoporosis Treatment Study Groups. Prevention of nonvertebral fractures by alendronate. *JAMA* 1997;277:1159-1164. (Class M)

Kelly TL, Crane G, Baran DT. Single X-ray absorptiometry of the forearm: precision, correlation, and reference data. *Calcif Tissue Int* 1994;54:212-218. (Class D)

Khovidhunkit W, Shoback DM. Clinical effects of raloxifene hydrochloride in women. *Ann Intern Med* 1999;130:431-439. (Class R)

Kleerekoper M, Edelson GW. Biochemical studies in the evaluation and management of osteoporosis: current status and future prospects. *Endoc Pract* 1996;2:13-19. (Class R)

Kohrt WM, Birge SJ. Differential effects of estrogen treatment on bone mineral density of the spine, hip, wrist and total body in late postmenopausal women. *Osteoporosis Int* 1995;5:150-155. (Class C)

Lang T, Augat P, Majumdar S, Ouyang X, Genant HK. Noninvasive assessment of bone density and structure using computed tomography and magnetic resonance. *Bone* 1998;22:149S-153S. (Class R)

Law MR, Wald NJ, Meade TW. Strategies for prevention of osteoporosis and hip fracture. *BMJ* 1991;303:453-459. (Class R)

Liberman UA, Weiss SR, Bröll J. Effect of oral alendronate on bone mineral density and the incidence of fractures in postmenopausal osteoporosis. *N Engl J Med* 1995;333:1437-1443. (Class A)

Lindsay R, Cosman F, Lobo RA, et al. Addition of alendronate to ongoing hormone replacement therapy in the treatment of osteoporosis: a randomized, controlled clinical trial. *J Clin Endocrinol Metab* 1999;84:3076-3081. (Class A)

Link TM, Majumdar A, Augat P, et al. Proximal femur: assessment for osteoporosis with T2*decay characteristics at MR imaging. *Radiology* 1998;209:531-536. (Class C)

Looker AC, Johnston CC, Wahner HW, et al. Prevalence of low femoral bone density in older U.S. women from NHANES III. *J Bone Miner Res* 1995a;10:796-802. (Class C)

Looker AC, Wahner HW, Dunn WL, et al. Proximal femur bone mineral levels of U.S. Adults. *Osteoporosis Int* 1995b;5:389-409. (Class C)

Lufkin EG, Whitaker MD, Nickelsen T, et al. Treatment of established postmenopausal osteoporosis with raloxifene: a randomized trial. *J Bone Miner Res* 1998;13:1747-1754. (Class A)

Madhok R, Melton LJ III, Atkinson EJ, O'Fallon WM, Lewallen DG. Urban vs rural increase in hip fracture incidence: age and sex of 901 cases 1980-89 in Olmsted County, U.S.A. *Acta Orthop Scand* 1993;64:543-548. (Class C)

Masud T, Jawed S, Doyle DV, Spector TD. A population study of the screening potential of assessment of trabecular pattern of the femoral neck (Singh index): the Chingford Study. *Br J Radiol* 1995(Apr);68(808):389-93. (Class C)

Matkovic V, Jelic T, Wardlaw G, et al. Timing of peak bone mass in Caucasian females and its implication for the prevention of osteoporosis. *J Clin Invest* 1994;93:799-808. (Class D)

Melton LJ III. Perspectives: How many women have osteoporosis now? *J Bone Miner Res* 1995;10:175-177. (Class R)

Melton LJ III, Atkinson EJ, O'Fallon WM, Wahner HW, Riggs BL. Long-term fracture prediction by bone mineral assessed at different skeletal sites. *J Bone Miner Res* 1993;8:1227-1233. (Class C)

Melton LJ III, Eddy DM, Johnston CC. Screening for osteoporosis. *Ann Intern Med* 1990;112:516-528. (Class R)

Michaëlsson K, Baron JA, Farahmand BY, et al. For the Swedish Hip Fracture Study Group. Hormone replacement therapy and risk of hip fracture: population based case-control study. *BMJ* 1998;316:1858-1863. (Class C)

Miller PD, Bonnick SL, Johnston CC, et al. The challenges of peripheral bone density testing. Which patients need additional central density skeletal measurements? *J Clin Densitometry* 1998;1:211-217. (Class R)

Mole PA, McMurdo ME, Paterson CR. Evaluation of peripheral dual energy X-ray absorptiometry: comparison with single photon absorptiometry of the forearm and dual energy X-ray absorptiometry of the spine or femur. *Br J Radiol* 1998;71:427-32. (Class C)

National Osteoporosis Foundation. Osteoporosis: physician's guide to prevention and treatment of osteoporosis. Washington, DC: National Osteoporosis Foundation, 1998a. (Class R)

National Osteoporosis Foundation. Osteoporosis: review of the evidence for prevention, diagnosis, and treatment and cost-effectiveness analysis. *Osteoporosis Int* 1998b(Suppl 4):S1-S88. (Class R)

Njeh CF, Boivin CM, Langton CM. The role of ultrasound in the assessment of osteoporosis: a review. *Osteoporosis Int* 1997;7:7-22. (Class R)

Office of Medical Applications of Research Consensus Conference. Osteoporosis. *JAMA* 1984;252:799-802. (Class R)

Ott S. Should women get screening bone mass measurements? *Ann Intern Med* 1986;104:874-876. (Class X)

Overgaard K, Hansen MA, Jensen SB, Christiansen C. Effect of salcatonin given intranasally on bone mass and fracture rates in established osteoporosis: a dose-response study. *BMJ* 1992;305:556-561. (Class A)

Owen RA, Melton LJ III, Johnson KA, Ilstrup DM, Riggs BL. Incidence of Colles' fracture in a North American community. *Am J Public Health* 1982;72:605-607. (Class C)

Patel R, Seah M, Blake GM, Jefferies AL, Crane FM, Fogelman I. Concordance and precision of dual X-ray absorptiometry with a 10 s scan. *Br J Radiol* 1996;69:816-20. (Class C)

Poor G, Jacobsen SJ, Melton LJ III. Mortality following hip fracture. In: Albarede JL, Garry PJ, Vellas P (eds.) *Facts and Research in Gerontology, 1994*. New York: Springer Publishing Co. 1994:91-109. (Class R)

Ray NF, Chan JK, Thamer M, Melton LJ III. Medical expenditures for the treatment of osteoporotic fractures in the United States in 1995: report from the National Osteoporosis Foundation. *J Bone Miner Res* 1997;12:24-35. (Class C)

Repa-Eschen LM. Prevention and treatment of osteoporosis in our healthcare delivery system. *The Osteoporotic Syndrome*. New York: Wiley-Liss. 1993:191-207. (Class R)

Reginster JY, Meurmans L, Zegels B, et al. The effect of sodium monofluorophosphate plus calcium on vertebral fracture rate in postmenopausal women with moderate osteoporosis. *Ann Intern Med* 1998;129:1-8. (Class A)

Resnick NM, Greenspan SL. 'Senile' osteoporosis reconsidered. *JAMA* 1989;261:1025-1029. (Class R)

Ribot C, Tremollieres F, Pouilles JM. Can we detect women with low bone mass using clinical risk factors? *Am J Med*1995;98(2A):52S-55S. (Class R)

Riggs BL, Melton LJ III. Involutional osteoporosis. *N Engl J Med* 1986;314:1676-1686. (Class R)

Riggs BL, Melton LJ III. The prevention and treatment of osteoporosis. *N Engl J Med* 1992;327:620-626. (Class R)

Riggs BL, Melton LJ III, O'Fallon WM. Drug therapy for vertebral fractures in osteoporosis: evidence that decrease in bone turnover and increases in bone mass both determine antifracture efficacy. *Bone* 1996;18:197S-201S. (Class R)

Riis BJ. Biochemical markers of bone turnover II: Diagnosis, prophylaxis, and treatment of osteoporosis. *Am J. Med* 1993;95(suppl 5A):17S-21S. (Class R)

Ross P, Huang C, Davis J et al. Predicting vertebral deformity using bone densitometry at various skeletal sites and calcaneus ultrasound. *Bone* 1995;16:325-332. (Class D)

Ross PD, Norimatsu H, Davis JW, et al. A comparison of hip fracture incidence among native Japanese, Japanese Americans, and American Caucasians. *Am J Epidemiol* 1991;133:801-809. (Class C)

Rubin SM, Cummings SR. Results of bone densitometry affect women's decisions about taking measures to prevent fractures. *Ann Intern Med* 1992;116:990-995. (Class D)

Rüegsegger P, Durand EP, Dambacher MA. Differential effects of aging and disease on trabecular and compact bone density of the radius. *Bone* 1991;12:99-105. (Class C)

Russell RG. The assessment of bone metabolism *in vivo* using biochemical approaches. *Horm Metab Res* 1997;29:138-144. (Class R)

Sartoris DJ, André M, Resnick C, Resnick D. Trabecular bone density in the proximal femur: quantitative CT assessment. *Radiology* 1986;160:707-712. (Class D)

Scientific Advisory Board of the National Osteoporosis Foundation. Clinical indications for bone mass measurements. *J Bone Miner Res* 1989;4(suppl 2):1-28. (Class R)

Sowers MF, Clark MK, Hollis B, Wallace RB, Jannausch M. Radial bone mineral density in pre- and perimenopausal women: a prospective study of rates and risk factors for loss. *J Bone Miner Res* 1992;7:647-657. (Class C)

Thompson P, Taylor J, Oliver R, Fisher A. Quantitative ultrasound (QUS) of the heel predicts wrist and osteoporosis-related fractures in women age 45-75 years. *J Clin Densitometry* 1998;1:219-225. (Class B)

van Daele PLA, Burger H, De Laet CEDH, Pols HAP. Ultrasound measurement of bone. *Clin Endocrinol* 1996;44:363-369. (Class R)

van Kuijk C, Genant HK. Detection of osteopenia and osteoporosis. In: Lobo RA (ed.) *Treatment of the Postmenopausal Woman: Basic and Clinical Aspects*. New York: Raven Press. 1994:169-174. (Class R)

Wimalawansa SJ. A four-year randomized controlled trial of hormone replacement and bisphosphonate, alone, or in combination, in women with postmenopausal osteoporosis. *Am J Med* 1998;104:219-226. (Class A)

Wimalawansa SJ. Combined therapy with estrogen and etidronate has an additive effect on bone mineral density in the hip and vertebrae: four-year randomized study. *Am J Med* 1995;99:36-42. (Class A)

Wasnich R. Bone mass measurement: prediction of risk. *Am J Med* 1993;95(suppl 5A):6S-10S. (Class R)

Whitaker MD. Clinical utility of biochemical markers of bone metabolism. *Endocr Pract* 1996;2:25-27. (Class R)

Yates AJ, Ross PD, Lydick E, Epstein RS. Radiographic absorptiometry in the diagnosis of osteoporosis. *Am J Med* 1995;98(2A):41S-47S. (Class R)

Appendix

See next page

Conclusion Grading Worksheet

Work Group's Conclusion: Quantitative ultrasound (QUS) can be used to predict hip fractures in women 65 years of age and older.

Conclusion Grade: II

Author/Year	Design Type	Class	Quality +,−,ø	Population Studied/Sample Size	Primary Outcome Measure(s)/Results (e.g., p-value, confidence interval, relative risk, odds ratio, likelihood ratio, number needed to treat)	Authors' Conclusions/ *Work Group's Comments (italicized)*
Hans et al. (1996)	Prospective Cohort	B	+	-5,662 women, ≥75 years (from EPIDOS cohort) -Excluded: history of hip fracture, Paget's disease, malignant bone disease, renal failure, hyperthyroidism, or treated hypothyroidism -Baseline exam, QUS of calcaneus, bone mineral density of femoral neck -Follow-up contact every 4 months for fracture assessment (with medical record verification)	-Mean follow-up of 2 years -115 first hip fractures (after minor trauma) -248 deaths from unrelated causes -Adjusted (age, weight, clinic site) relative risk of hip fracture with one SD reduction: BUA 2.0 (95%CI: 1.6-2.4) SOS 1.7 (95%CI: 1.4-2.1) BMD 1.9 (95%CI: 1.6-2.4) -Rate of hip fracture (per 1000 woman-years) based on BMD and BUA: BUA above median, BMD above median=2.7 BUA below median, BMD above median=6.2 BUA above median, BMD below median=11.8 BUA below median, BMD below median=19.6	-Ultrasonographic measurements at the calcaneus predict the risk of hip fracture as accurately as does femoral neck bone mineral density. NOTES: follow-up compliance was 97.7% (2.2% refused to continue to participate and 0.1% were lost)
Bauer et al. (1997)	Prospective cohort	B	+	-6,189 women >65 years (from SOF cohort that had been identified 5 years earlier) -BUA of heel; BMD of hip, calcaneus -Follow-up contact every 4 months for fracture assessment (verified) -Excluded spine fractures	-Mean follow-up of 2 years -350 fractures (54 hip, 70 wrist, 40 humerus, 36 ankle); those with fractures were older and had lower BUA and BMD ($p<0.001$) -Adjusted (age, clinic site) relative risk (95% CI) of fracture with one SD reduction: BUA / BMDc* / BMDf* Non-spine 1.3 (1.2-1.5) 1.4 (1.2-1.6) 1.3 (1.1-1.5) Hip 2.0 (1.5-2.7) 2.2 (1.9-3.0) 2.6 (1.9-3.8) Fem. Neck 1.3 (0.9-2.0) 1.4 (0.9-2.1) 2.0 (1.2-3.1) Trochanter 3.3 (2.0-5.5) 3.4 (2.1-5.1) 3.9 (2.3-6.8) *BMDc=calcaneal BMD, BMDf=femoral neck BMD -After adjustment for hip BMD, low BUA measurement was still associated with increased risk for non-spine fracture (RR=1.2; 95% CI=1.1-1.4 or any hip fracture RR=1.5; 95%CI=1.0-2.1) -Relative risk of hip fracture among women in highest quartile of predicted risk vs. 3 lower quartiles: BUA only 2.5 (1.3-4.9) BMDf only 4.5 (2.4-8.4) BUA and BMDf 5.0 (2.7-9.5)	-BUA is a strong predictor of hip and all non-spine fractures with a strength of association similar to that observed with BMD NOTES: the women included in this study represent 72% of the original cohort subjects who were known to be alive; ultrasound velocity was not available

Author/Year	Design Type	Class	Quality +,–,ø	Population Studied/Sample Size	Primary Outcome Measure(s)/Results (e.g., p-value, confidence interval, relative risk, odds ratio, likelihood ratio, number needed to treat)	Authors' Conclusions/ *Work Group's Comments (italicized)*
Huang et al., 1998	Prospective Cohort	B	–	-560 post-menopausal women ages 55 to 92 years (from the Hawaii Osteoporosis Study) -BUA of the heel -Fractures (non-spine, non-violent) identified by self-report (verified by medical records)	-Mean follow-up time of 2.7 years -66 women with at least one fracture -Adjusted (age) odds ratios (95%CI) for fracture with one SD reduction in BUA: Vertebral fracture (39 cases) 1.50 (1.05-2.16) Non-spine fractures (31 cases) 1.89 (1.27-2.88) Any fracture (66 cases) 1.72 (1.30-2.31)	-Calcaneal BUA is a significant predictor of non-spine fracture, vertebral fracture, and overall fracture risk. NOTES: odds ratios are based on a relatively small number of fracture cases observed within a relatively short period of follow-up *Work Groups' Comments:* *-It was not specified how the fracture data was obtained from the women*
Thompson et al., 1998	Prospective Cohort	B	ø	-3,143 women ages 45 to 75 (peri- and post-menopausal) -QUS of heel (calculated BUA, SOS, and stiffness index) -Fractures identified from practice records	-Mean follow-up time of 31 months -150 fractures (7 women had 2 fractures which were counted as separate incidents) -Adjusted (age) odds ratios (95%CI) for fracture with one SD change in QUS parameter: Fracture (#) BUA SOS Stiffness All (150) 1.4 (1.1-1.6) 1.3 (1.1-1.5) 1.5 (1.3-1.8) Wrist (63) 1.6 (1.2-2.1) 1.5 (1.2-2.1) 1.8 (1.4-2.4) Osteo. (26)* 1.9 (1.3-2.9) 1.6 (1.0-2.5) 2.2 (1.4-3.4) Other (61) 1.0 (0.8-1.3) 1.1 (0.8-1.3) 1.1 (0.9-1.5) *Osteoporosis-related fractures (proximal humerus, hip, pelvis, and vertebrae) -Data were divided by age groups (46-55, 56-65, 66-75); BUA was a significant predictor of all fractures, osteoporotic fractures, and wrist fracture except in the 46-55 age group; SOS was a significant predictor of osteoporotic fractures in the 66-75 age group	-QUS may be used to predict osteoporosis-related fractures and particularly wrist fractures in postmenopausal women over the age of 55 years as well as hip fractures in older women NOTES: 37 women were lost to follow-up; few reported vertebrae fractures (likely because they required documented evidence from a hospital that a fracture had occurred); did not validate diagnosis of fracture with hospital records or x-rays; overall there were few fractures in the 46-55 age group