chapter 15

Image Quality in Nuclear Medicine

Image quality refers to the faithfulness with which an image represents the imaged object. The quality of nuclear medicine images is limited by several factors. Some of these factors, relating to performance limitations of the gamma camera, already have been discussed in Chapter 14. In this chapter, we discuss the essential elements of image quality in nuclear medicine and how it is measured and characterized. Because of its predominant role in nuclear medicine, the discussion focuses on planar imaging with the gamma camera; however, the general concepts are applicable as well to the tomographic imaging techniques that are discussed in Chapters 17 to 19.

A. BASIC METHODS FOR CHARACTERIZING AND EVALUATING IMAGE QUALITY

There are two basic methods for characterizing or evaluating image quality. The first is by means of *physical characteristics* that can be quantitatively measured or calculated for the image or imaging system. Three such characteristics that are used for nuclear medicine image quality are (1) *spatial resolution* (detail or sharpness), (2) *contrast* (difference in image density or intensity between areas of the imaged object containing different concentrations of radioactivity), and (3) *noise* (statistical noise caused by random fluctuations in radioactive decay, or structured noise, e.g., resulting from instrument artifacts). Although they describe three different aspects of image quality, these three factors cannot be treated as completely independent parameters because improvements in one of them frequently are obtained at the expense or deterioration of one or more of the others. For example, improved collimator resolution usually involves a tradeoff of decreased collimator efficiency (see Chapter 14, Section C) and, hence, decreased counting rates and increased image statistical noise.

The second method for characterizing or evaluating image quality is by means of *human observer performance studies* using images obtained with different imaging systems or under different imaging conditions. Although observer performance can be characterized objectively, and certainly is related to the physical measures of image quality described earlier, the relationships are not well established because of the complexity of the human visual system and other complicating factors, such as observer experience. Hence, the two methods, though related, are somewhat independent.

A related approach, known as *computer observer performance studies,* uses a mathematical model that under appropriate conditions predicts the performance of a human observer and can be used as a surrogate for actual human observer studies. Because human observer studies require large numbers of images and therefore are very time consuming, computer observers often are more practical. Details regarding computer observer models are beyond the scope of this text and the interested reader is referred to reference 1 for further information.

B. SPATIAL RESOLUTION

1. Factors Affecting Spatial Resolution

Spatial resolution refers to the sharpness or detail of the image, or to the ability of the imaging instrument to provide such sharpness or detail. The sample images presented in Chapters 13 and 14 already have demonstrated that nuclear medicine images have somewhat limited spatial resolution, at least

in comparison with photographic or radiographic images. A number of factors contribute to the lack of sharpness in these images.

Collimator resolution is perhaps the principal limiting factor when absorptive collimators are used for spatial localization (Chapter 14, Section C). Because collimator hole diameters must be relatively large (to obtain reasonable collimator efficiency), there is blurring of the image by an amount at least as great as the hole diameters (Equation 14-6). Collimator resolution also depends on source-to-detector distance (Figs. 14-16 through 14-22). Note that collimator resolution is not a factor in positron emission tomography (PET) imaging, which uses annihilation coincidence detection for spatial localization (Chapter 18, Section A).

A second factor is *intrinsic resolution* of the imaging detector. With the gamma camera, this limitation arises primarily because of statistical variations in the distribution of light photons among the photomultiplier tubes (Chapter 14, Section A.1). Intrinsic resolution is a function of γ-ray energy with the gamma camera, becoming poorer with decreasing γ-ray energy (Fig. 14-2). For imaging devices with discrete detector elements, such as many PET systems (see Chapter 18, Section A.3), the size of the individual detector elements largely determines the intrinsic resolution of the device.

Image sharpness also can be affected by patient motion. Figure 15-1 shows images of a brain phantom obtained with and without motion. Respiratory and cardiac motion can be especially troublesome because of the lengthy imaging times required in nuclear medicine and the relatively great excursions in distance (2-3 cm) that are possible in these instances. Gated-imaging techniques (see Chapter 20, Section A.4) have been employed to minimize motion blurring, especially in cardiac studies. Breath-holding also has been used to minimize blurring caused by respiratory motion.

Nuclear medicine imaging systems acquire data on a discrete matrix of locations, or pixels, which leads to *pixelation effects* in the image. As discussed in Chapter 20, the size of the discrete pixels sets a limit on the spatial resolution of the image. In general, it is desirable to have at least two pixels per full width at half maximum (FWHM) of system resolution to avoid creating distracting pixelation effects and possible loss of image detail.

2. Methods for Evaluating Spatial Resolution

Spatial resolution may be evaluated by subjective or objective means. A subjective evaluation can be obtained by visual inspection of images of organ phantoms that are meant to simulate clinical images (e.g., the brain phantom in Fig. 15-1). Although they attempt to project "what the physician wants to see," organ phantoms are not useful for quantitative comparisons of resolution between different imaging systems or techniques. Also, because of the subjective nature of the evaluation, different observers might give different interpretations of comparative image quality.

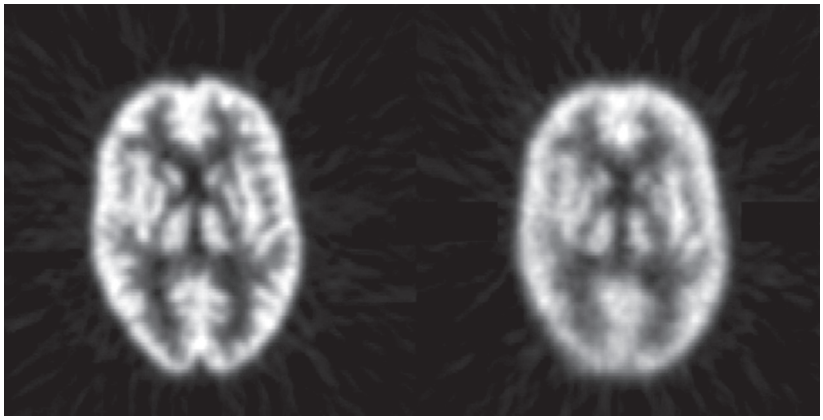

FIGURE 15-1 Images of a brain phantom obtained with phantom stationary (*left*) and with random translations (several mm) and rotations (several degrees) (*right*) during the imaging procedure, demonstrating motion-blurring effects. (*Adapted from Fulton R et al: Accuracy of motion correction methods for PET brain imaging. 2004 IEEE Nuclear Science Symposium Conference Record, 4226-4230.*)

A phantom that can be used for more objective testing of spatial resolution is shown in Figure 15-2. *Bar phantoms* are constructed of lead or tungsten strips, generally encased in a plastic holder. Strips having widths equal to the spaces between them are used. For example, a "5-mm bar pattern" consists of 5-mm-wide strips separated edge to edge by 5-mm spaces. The four-quadrant bar phantom shown in Figure 15-2 has four different strip widths and spacings. To evaluate the intrinsic resolution of a gamma camera, the bar phantom is placed directly on the uncollimated detector and irradiated with a uniform radiation field, typically a point source of radioactivity at several meters distance from the detector. To evaluate the resolution with a collimator, the phantom is placed directly on the collimated detector and irradiated with a point source at several meters distance, or with a sheet source of radioactivity placed directly behind the bar phantom. Spatial resolution is expressed in terms of the smallest bar pattern visible on the image. There is a certain amount of subjectivity to the evaluation, but not so much as with organ phantoms.

To properly evaluate spatial resolution with bar phantoms, one must ensure that the thickness of lead strips is sufficient so that they are virtually opaque to the γ rays being imaged. Otherwise, poor visualization may be due to poor contrast of the test image rather than poor spatial resolution of the imaging device. For ^{99m}Tc (140 keV) and similar low-energy γ-ray emitters, tenth-value thicknesses in lead are approximately 1 mm or less, whereas for ^{131}I (364 keV), annihilation photons (511 keV), and so on, they are on the order of 1 cm (see Table 6-4). Most commercially available bar phantoms are designed for ^{99m}Tc and are not suitable for use with higher-energy γ-ray emitters.

A still more quantitative approach to evaluating spatial resolution is by means of the point-spread function (PSF) or line-spread function (LSF). General methods for recording these functions were described in Chapter 14, Section E.2. Examples of LSFs are shown in Figure 17-8 for a single-photon emission computed tomography (SPECT) camera and in Figure 18-5 for a PET system. Although the complete profile is needed to fully characterize spatial resolution, a partial specification is provided by its FWHM (Fig. 14-15). The FWHM is not a complete specification because PSFs or LSFs of different shapes can have the same FWHM. (Compare, for example, the different shapes in Figs. 18-5 and 18-7). However, the FWHM is useful for general comparisons of imaging devices and techniques. Roughly speaking, the FWHM of the PSF or LSF of an imaging instrument is approximately 1.4-2 times the width of the smallest resolvable bar pattern (Fig. 15-3). Thus an instrument having an FWHM of 1 cm should be able to resolve 5- to 7-mm bar patterns.

In most cases, multiple factors contribute to spatial resolution and image blurring. The method for combining FWHMs for intrinsic and collimator resolutions to obtain the overall system FWHM is discussed in Chapter 14 , Section C.4 and in Appendix G. In general, if a system has n factors or components that each contribute independently to blurring, individually characterized by FWHM_1, FWHM_2, . . ., FWHM_n, the FWHM for the system is given by

$$\text{FWHM}_{\text{sys}} \approx \sqrt{\text{FWHM}_1^2 + \text{FWHM}_2^2 + \ldots + \text{FWHM}_n^2} \quad (15\text{-}1)$$

This equation provides an exact result when all of the components have gaussian-shaped

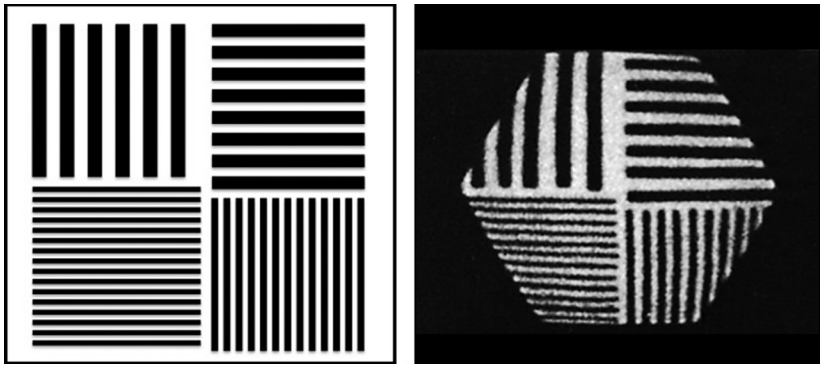

FIGURE 15-2 Design (*left*) and gamma camera image (*right*) of a four-quadrant bar phantom used for evaluation of spatial resolution.

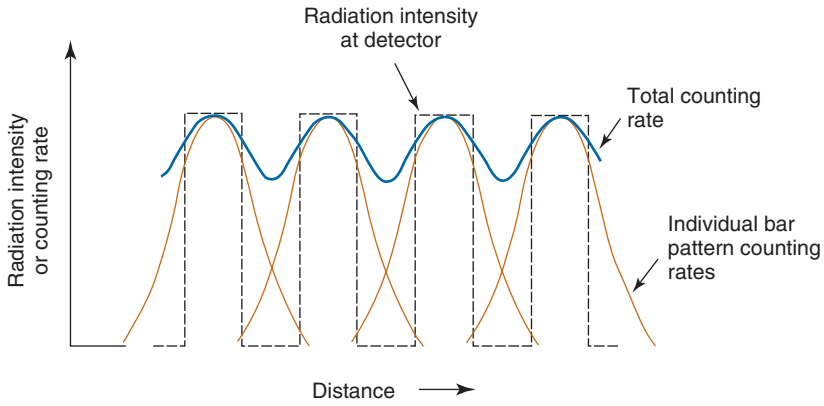

FIGURE 15-3 Counting-rate profiles obtained on a bar pattern phantom with an imaging system having FWHM resolution approximately 1.6 times the width of individual bars and spaces.

blurring functions, but it is an approximation when nongaussian shapes are involved. Note that if the FWHM for any one factor is significantly larger than the others, it becomes the dominating factor for system FWHM. Thus, for example, if $FWHM_1 \gg FWHM_2$, it makes little sense to expend substantial effort toward improving $FWHM_2$.

The most detailed specification of spatial resolution is provided by the *modulation transfer function* (MTF). The MTF is the imaging analog of the frequency response curve used for evaluating audio equipment. In audio equipment evaluations, pure tones of various frequencies are fed to the input of the amplifier or other component to be tested, and the relative amplitude of the output signal is recorded. A graph of relative output amplitude versus frequency is the frequency response curve for that component (Fig. 15-4). A system with a "flat" curve from lowest to highest frequencies provides the most faithful sound reproduction.

By analogy, one could evaluate the fidelity of an imaging system by replacing the audio tone with a "sine-wave" distribution of activity (Fig. 15-5). Instead of varying in time (cycles per second), the activity distribution varies with distance (cycles per centimeter or cycles per millimeter). This is called the *spatial frequency* of the test pattern, customarily symbolized by k.* The modulation of the test pattern, which is a measure of its contrast, is defined by

$$M_{in} = (I_{max} - I_{min})/(I_{max} + I_{min}) \quad (15\text{-}2)$$

where I_{max} and I_{min} are the maximum and minimum radiation intensities emitted by the test pattern. M_{in} is the input modulation for the test pattern and ranges from zero ($I_{max} = I_{min}$, no contrast) to unity ($I_{min} = 0$, maximum contrast). Similarly, output modulation M_{out} is defined in terms of the modulation of output image (e.g., image density or counting rate recorded from the test pattern).

$$M_{out} = (O_{max} - O_{min})/(O_{max} + O_{min}) \quad (15\text{-}3)$$

*Technically speaking, the notation k is used in physics to denote "cycles per radian," and the notation $\bar{k}$ or "k-bar" is used to denote "cycles per distance." Mathematically, $\bar{k} = k/2\pi$, because there are 2π radians per cycle. For notational simplicity, we use k for cycles per distance in this text.

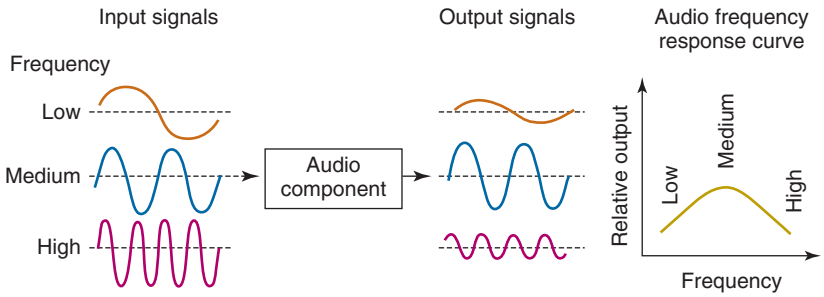

FIGURE 15-4 Basic principles for generating frequency response curves for an audio system.

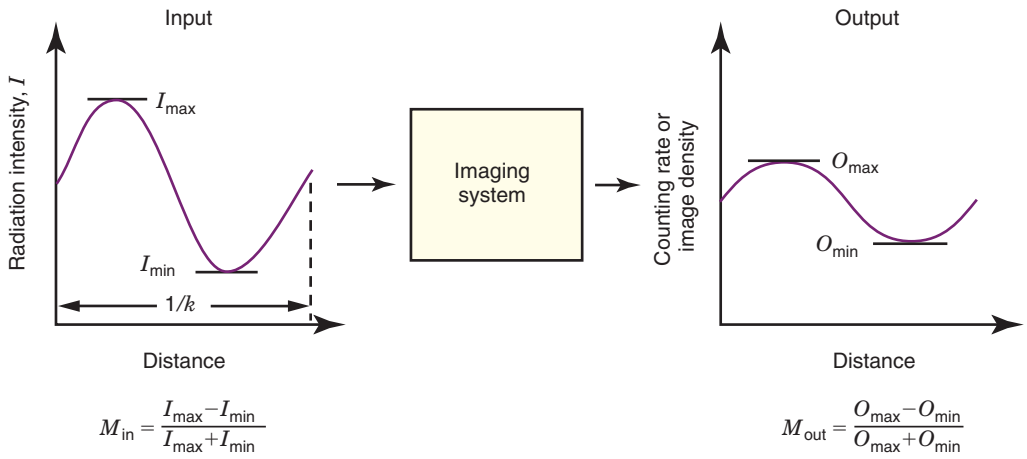

FIGURE 15-5 Basic principles for determining the modulation transfer function of an imaging instrument. Input contrast is measured in terms of object radioactivity or emission rate. Output contrast is measured in terms of counting rate, image intensity, etc. Spatial frequency is k.

The ratio of output to input modulation is the MTF for the spatial frequency k of the test pattern,

$$\text{MTF}(k) = M_{\text{out}}(k)/M_{\text{in}}(k) \qquad (15\text{-}4)$$

The usefulness of the MTF (or frequency response curve) derives from the fact that any image (or audio signal) can be described as a summation of sine waves of different frequencies. For audio signals, the sound "pitch" is determined by its basic sine-wave frequency, whereas superimposed higher frequencies create the unique sound characteristics of the instrument or human voice producing it. An audio system with a flat frequency response curve over a wide frequency range generates an output that matches faithfully the sound of the instrument or voice producing it. Inexpensive audio systems generally reproduce the midrange audio frequencies accurately but have poor response at low and high frequencies. Thus they have poor bass response (low frequencies) and poor sound "quality" (high frequencies).

An imaging system with a flat MTF curve having a value near unity produces an image that is a faithful reproduction of the imaged object. Good low-frequency response is needed to outline the coarse details of the image and is important for the presentation and detection of relatively large but low-contrast lesions. Good high-frequency response is necessary to portray fine details and sharp edges. This is of obvious importance for small objects but sometimes also for larger objects because of the importance of edges and sharp borders for detection of low-contrast objects and for accurate assessment of their size and shape.

Figure 15-6 illustrates some typical MTF curves for a gamma camera collimator. The MTF curves have values near unity for low frequencies but decrease rapidly to zero at higher frequencies. Thus the images of a radionuclide distribution obtained with this collimator show the coarser details of the distribution faithfully but not the fine details. Edge sharpness, which is a function of the high-frequency MTF values, also is degraded. This type of performance is characteristic of virtually all nuclear medicine imaging systems. Note also that the MTF curve at higher frequencies decreases more rapidly with increasing source-to-collimator distance.

The MTF curve characterizes completely and in a quantitative way the spatial resolution of an imaging system for both coarse and fine details. Images of bar patterns and similar test objects are quantitative only for specifying the limiting resolution of the imaging system, for example, the minimum resolvable bar pattern spacing. Bar-pattern images and MTF curves can be related semiquantitatively by noting that the spatial frequency of a bar pattern having bar widths and spaces of x cm is one cycle per $2x$ cm. Thus a "5-mm bar pattern" has a basic spatial frequency of one cycle per centimeter (one bar and one space per centimeter). Roughly speaking, bar patterns are no longer visible when the MTF for

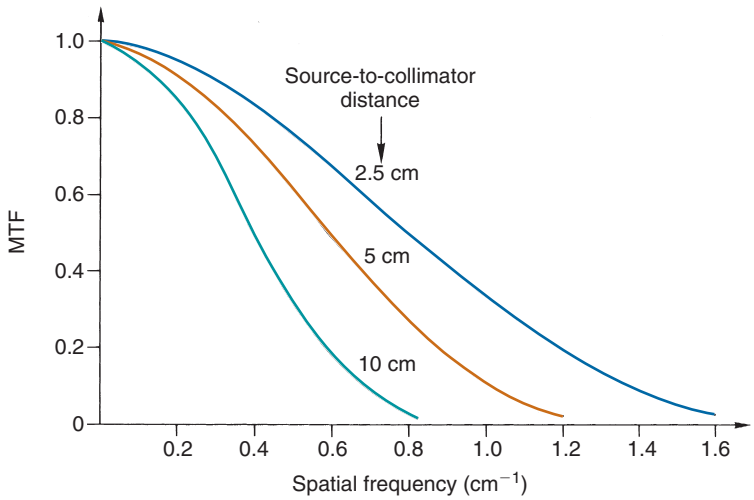

FIGURE 15-6 Modulation transfer function curves for a typical parallel-hole collimator for different source-to-collimator distances. (*Data from Ehrhardt JC, Oberly LW, Cuevas JM:* Imaging Ability of Collimators in Nuclear Medicine. *Publication No. [FDA] 79-8077. Rockville, MD, U.S. Department of Health, Education, and Welfare, 1978.*)

their basic spatial frequency drops below a value of approximately 0.1. MTF curves thus can be used to estimate the minimum resolvable bar pattern for an imaging system.

In practice, MTFs are not determined using sinusoidal activity distributions, as illustrated in Figure 15-5, which would be difficult to construct. Instead, they are obtained by mathematical analysis of the LSF or PSF. Specifically, the MTF of an imaging system can be derived from the Fourier transform (FT) of the LSF or PSF.* The one-dimensional (1-D) FT of the LSF is the MTF of the system measured in the direction of the profile, that is, perpendicular to the line source. Similarly, the 1-D FT of a profile recorded through the center of the PSF gives the MTF of the system in the direction of the profile. Alternatively, a 2-D FT of the 2-D PSF provides a 2-D MTF that can be used to determine the frequency response of the system at any angle relative to the imaging detector. This sometimes is useful for imaging systems that have asymmetrical spatial resolution characteristics, such as detector arrays with rectangular elements. Some PET detector arrays have this property (see Chapter 18, Section B).

It also is possible to obtain a 3-D representation of the MTF from a complete 3-D data set for the PSF. This is potentially useful for characterizing the spatial resolution of tomographic instruments in all three spatial directions. Note that in all cases the diameter or width of the source should be much smaller than resolution capability of the imaging device ($d \leq$ FWHM/4). Additional discussions about the measurement of MTFs and their properties can be found in references 2 and 3.

Another useful feature of MTFs is that they can be determined for different components of an imaging system and then combined to determine the system MTF. This feature allows one to predict the effects of the individual components of the system on the MTF of the total system. For example, one can obtain the MTF for the intrinsic resolution of the Anger camera detector, $\text{MTF}_{\text{int}}(k)$, and another for the collimator, $\text{MTF}_{\text{coll}}(k)$. The system MTF then is obtained by point-by-point multiplication of the intrinsic and collimator MTFs at each value of k:

$$\text{MTF}_{\text{sys}}(k) = \text{MTF}_{\text{int}}(k) \times \text{MTF}_{\text{coll}}(k) \quad (15\text{-}5)$$

In general, the MTF of a system is the product of the MTFs of its components.

If two systems have MTF curves of the same general shape, one can predict confidently that the system with the higher MTF values will have superior spatial resolution; however, the situation is more complicated when comparing two systems having MTF curves of different shapes. For example, Figure 15-7 shows MTF curves for two collimators, one of which would be better for visualizing large low-contrast structures (low frequencies), the other for fine details (high

*More specifically, the MTF is the modulus, or amplitude, of the FT, the latter generally being a complex number. See Appendix F for a more detailed discussion of FTs.

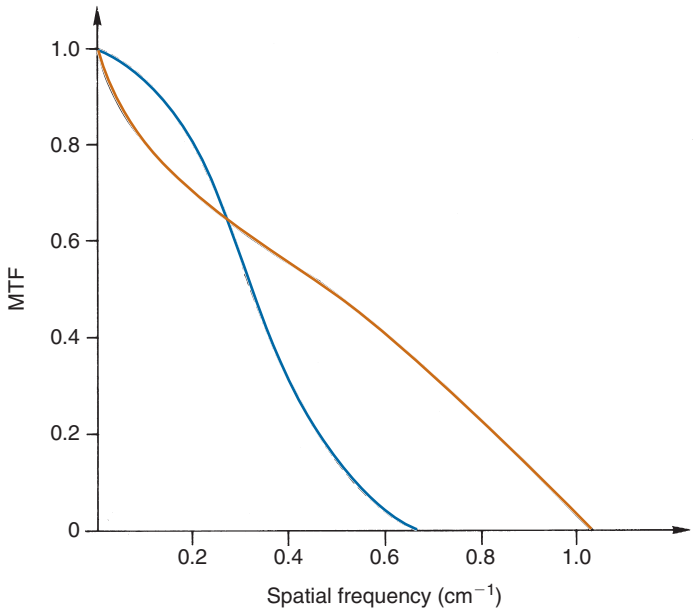

FIGURE 15-7 Modulation transfer function (MTF) curves for two different collimators. One has better low-frequency resolution for coarse details (*blue line*), whereas the other is better for fine details (*orange line*). (*Data from Ehrhardt JC, Oberly LW, Cuevas JM:* Imaging ability of collimators in nuclear medicine. *Rockville, MD, U.S. Department of Health, Education, and Welfare, Publ. No. [FDA] 79-8077, 1978, p 20.*)

frequencies). To gain an impression of comparative image quality in this situation, one would probably have to evaluate organ phantoms or actual patient images obtained with these collimators.

C. CONTRAST

Image contrast refers to differences in intensity in parts of the image corresponding to different levels of radioactive uptake in the patient. In nuclear medicine, a major component of image contrast is determined by the properties of the radiopharmaceutical. In general, it is desirable to use an agent having the highest lesion-to-background uptake or concentration ratio. Some aspects of radiopharmaceutical design that affect this issue were discussed in Chapter 5, Section F. Physical factors involved in image formation also can affect contrast. In general, factors that affect contrast in nuclear medicine also affect the statistical noise levels in the image. More specifically, they affect the contrast-to-noise ratio (CNR), which is discussed in detail in the next section. Here we focus only on some factors that affect contrast.

A general definition of contrast is that it is the ratio of signal change of an object of interest, such as a lesion, relative to the signal level in surrounding parts of the image. Thus if R_o is the counting rate over normal tissue and R_ℓ is the counting rate over a lesion, the contrast of the lesion is defined as

$$C_\ell = \frac{R_\ell - R_o}{R_o} = \frac{\Delta R_\ell}{R_o} \qquad (15\text{-}6)$$

where ΔR_ℓ is the change in counting rate over the lesion relative to the surrounding background.* Contrast sometimes is expressed as a percentage, for example, $C_\ell = 0.1$ = "10% contrast."

Perhaps the major factor affecting contrast is *added background counting rates* that are superimposed more or less uniformly over the activity distribution of interest. For example, suppose that in the absence of background counts a certain object (e.g., a lesion) has

*This equation is related to, but not the same as, the equations for modulation given in Equations 15-2 and 15-3. The definition used here has the disadvantage that it does not apply when $R_o = 0$. However, this situation rarely, if ever, applies in nuclear medicine, and the definition in Equation 15-6 is more straightforward for the analysis of contrast and CNR.

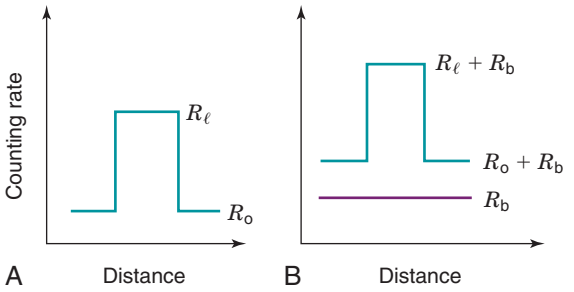

FIGURE 15-8 Effect on image contrast of adding a background counting rate R_b.

intrinsic contrast as defined by Equation 15-6. Suppose then that a uniform background counting rate R_b is superimposed on the image (Fig. 15-8). Then the lesion contrast becomes

$$\begin{aligned} C'_\ell &= \frac{(R_\ell + R_b) - (R_o + R_b)}{R_o + R_b} \\ &= \frac{\Delta R_\ell}{R_o + R_b} \\ &= \frac{\Delta R_\ell}{R_o} \times \left[\frac{1}{1 + (R_b / R_o)}\right] \\ &= C_\ell \times \left[\frac{1}{1 + (R_b / R_o)}\right] \end{aligned} \qquad (15\text{-}7)$$

Comparing the last line of Equation 15-7 with Equation 15-6, it can be seen that contrast is decreased by the additional factor R_b/R_o in the denominator.

EXAMPLE 15-1

Suppose that under ideal conditions, a certain radiopharmaceutical produces lesion and normal tissue counting rates given by R_o and $R_\ell = 1.2R_o$, respectively. Suppose further that a background counting rate $R_b = R_o$ then is added to the image. Calculate the image contrast with and without the added background counting rate.

Answer

Using Equation 15-6 for the intrinsic contrast without the added background

$$\begin{aligned} C_\ell &= \frac{(1.2R_o - R_o)}{R_o} \\ &= 0.2\ (20\%) \end{aligned}$$

When background amounting to $R_b = R_o$ is added, according to Equation 15-7

$$\begin{aligned} C'_\ell &= 0.2 \times \left[\frac{1}{1 + (R_b / R_o)}\right] \\ &= 0.2 \times \left[\frac{1}{1+1}\right] \\ &= 0.2 \times \frac{1}{2} = 0.1\ (10\%) \end{aligned}$$

Thus contrast is reduced by 50% by the added background.

Example 15-1 illustrates that added background can reduce image contrast substantially. It should be noted again that background counting rates also add to the noise levels in the image, just as they add to the noise levels in counting measurements (see Chapter 9, Section D.4). This is discussed in more detail in Section D.

Background counting rates can arise from a number of sources. Septal penetration and scattered radiation are two examples. Another would be inadequately shielded radiation sources elsewhere in the imaging environment. Septal penetration is avoided by using a collimator that is appropriately designed for the radionuclide of interest (Chapter 14, Section C.2). Scattered radiation can be minimized by pulse-height analysis; however, sodium iodide [NaI(Tl)] systems cannot reject all scatter, and rejection becomes especially difficult for γ-ray energies below approximately 200 keV, as illustrated by Figure 10-10. Using a narrower analyzer window for scatter rejection also decreases the recorded counting rate and increases the statistical noise in the image. A reasonable tradeoff between counting efficiency and scatter rejection for imaging systems using NaI(Tl) detectors is obtained with a 15% energy window centered on the γ-ray photopeak. There has been continuing interest in applying semiconductor detectors to nuclear medicine imaging to take advantage of their superior energy resolution for discrimination against scattered radiation by pulse-height analysis (see Figs. 10-14 and 10-15).

Figure 15-9 shows the effect of scattered radiation on images of a phantom. With a very wide analyzer window, there is virtually no rejection of scattered radiation and a noticeable loss of image contrast. The loss of contrast can result in degraded visibility of both large low-contrast objects and fine details in the image. Figure 15-10, for example, illustrates the effects of scattered radiation (or

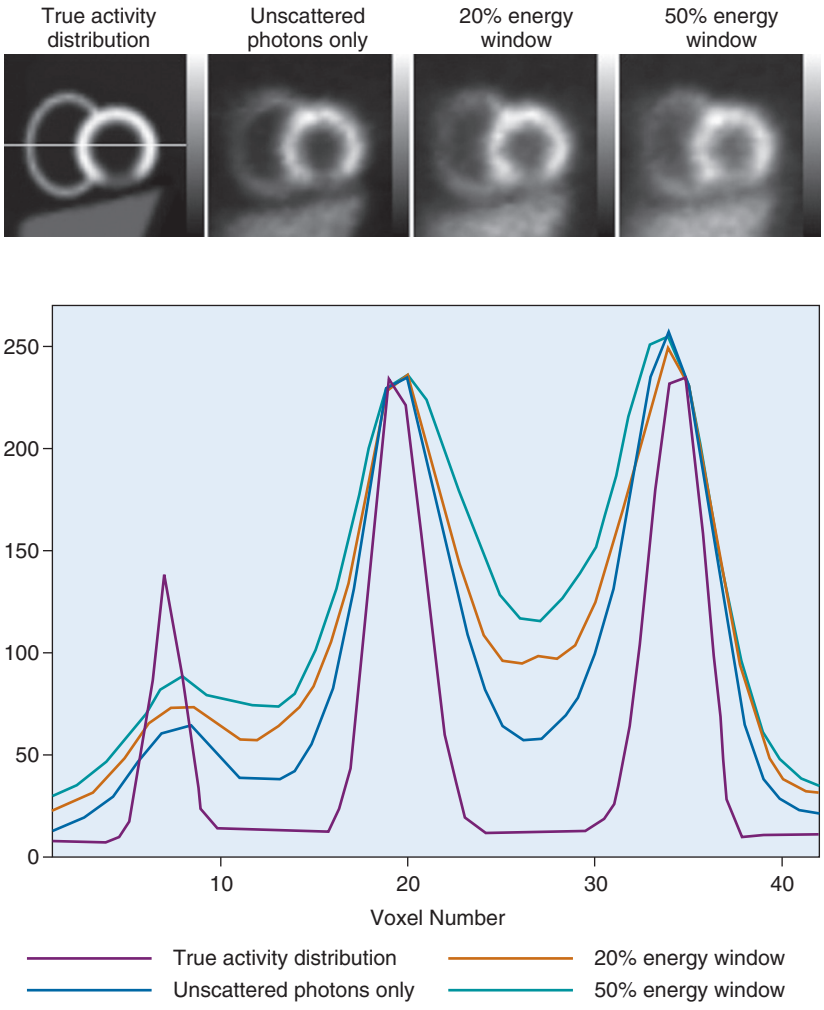

FIGURE 15-9 Effect of scatter and pulse-height analysis on image contrast. The images were generated by Monte Carlo simulations mimicking a clinical study of myocardial function using the radiotracer ^{201}TlCl. Count profiles through the images also are shown. These profiles are taken along the line shown in the image of the true activity distribution. The images also demonstrate blurring of the activity distribution caused by the finite camera resolution. (*Courtesy Dr. Hendrik Pretorius and Dr. Michael King, University of Massachusetts Medical School, Worcester, MA*).

septal penetration, which has similar effects) on the LSF and MTF of an imaging system. The addition of long "tails" to the LSF results first in the suppression of the MTF curve at low frequencies. This is reflected in poorer contrast of large objects that would make large low-contrast objects more difficult to detect or characterize. The high-frequency portion of the MTF curve also is suppressed, which has the effect of shifting the limiting frequency for detection of high-contrast objects (e.g., bar patterns) to lower frequencies. Thus the contrast-degrading effects of added background decrease the visibility of *all* structures in the image, particularly those that may already be near the borderline of detectability. These effects are apparent in Figure 15-9, which demonstrates a perceptible loss of image sharpness as well as overall image contrast when the added background is present.

An important contributor to background radiation in conventional planar imaging is radioactivity above and below the object of interest. Image contrast is improved in emission computed tomography (SPECT and PET) (see Chapters 17 and 18) because it permits imaging of an isolated slice without the superimposed activities in overlying and underlying structures. Tomographic techniques offer

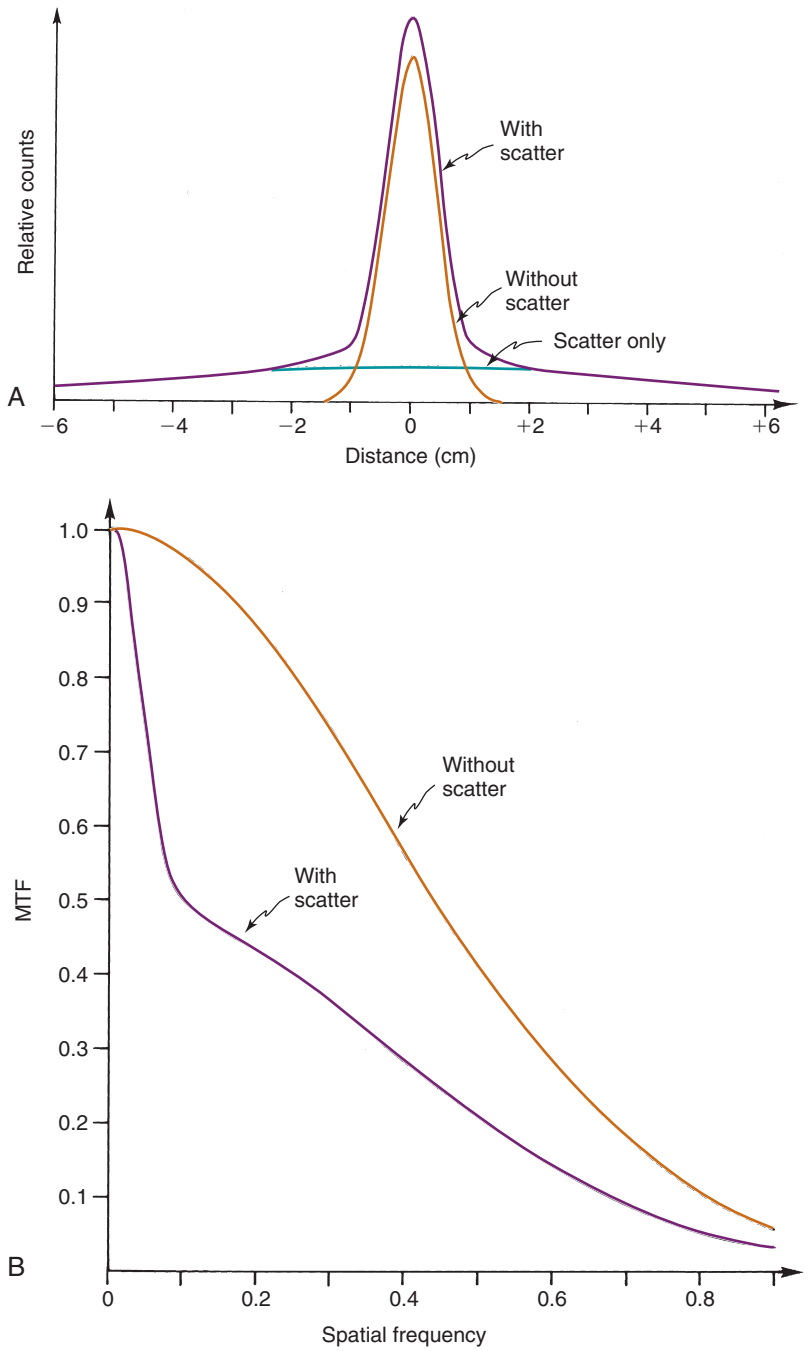

FIGURE 15-10 Illustration of effects of scatter and septal penetration on line-spread function (LSF) (*A*) and modulation transfer function (MTF) (*B*) of an imaging system. The long "tails" on the LSF have the effect of suppressing the MTF curve at both low and high spatial frequencies.

significant improvements for the detection of low-contrast lesions. Figure 15-11 illustrates this effect. Details of emission computed tomographic imaging are presented in Chapters 16 to 18; however, even at this point, the benefits of removing the interfering effects of overlying and underlying activity should be evident.

The preceding discussion relates to the effects of various types of background radiation on input contrast to the imaging system. It is possible with computers to apply "background subtraction" or "contrast enhancement" algorithms and thereby restore the original contrast, at least in terms of the

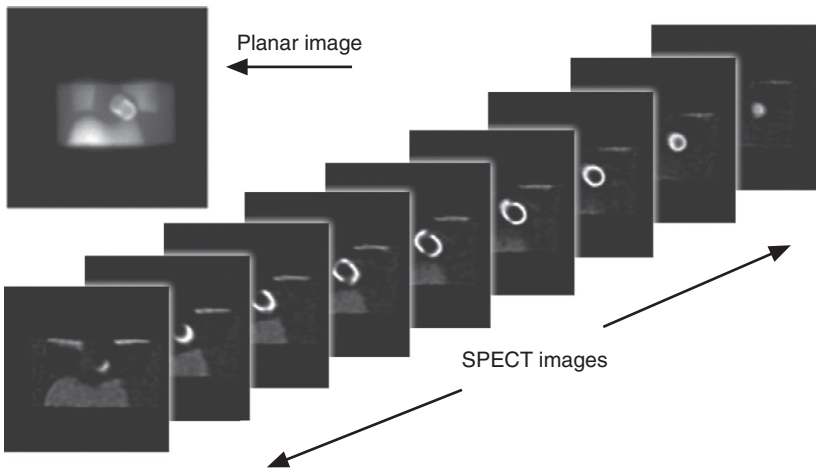

FIGURE 15-11 Planar (*upper left*) and single-photon emission computed tomographic (SPECT) (*center*) images of a thoracic phantom. Note the improved contrast and visibility of the voids in the cardiac portion of the phantom when overlying and underlying activity are removed in the SPECT images. (*Courtesy Dr. Freek Beekman, Delft University of Technology, Netherlands.*)

relative brightness levels between a lesion and its surrounding area. However, these techniques also enhance the statistical noise levels in the image as well as the contrast of any underlying artifacts, such as gamma camera image nonuniformities. Thus the critical parameter to consider regarding computer enhancement techniques is their effect on CNR. This concept is discussed in the following section.

D. NOISE

1. Types of Image Noise

Image noise generally can be characterized as either random or structured. *Random noise* refers to the mottled appearance of nuclear medicine images caused by random statistical variations in counting rate (Chapter 9). This is a very important factor in nuclear medicine imaging and is discussed in detail in this section.

Structured noise refers to nonrandom variations in counting rate that are superimposed on and interfere with perception of the object structures of interest. Some types of structured noise arise from the radionuclide distribution itself. For example, in planar imaging, uptake in the ribs may be superimposed over the image of the heart in studies to detect myocardial infarction with ^{99m}Tc-labeled pyrophosphates. Bowel uptake presents a type of structured noise in studies to detect inflammation or abscess with ^{67}Ga.

Structured noise also can arise from imaging system artifacts. Nonuniformities in gamma camera images (see Fig. 14-10) are one example. Various "ring" or "streak" artifacts generated during reconstruction tomography are another (e.g., see Figs. 16-11 and 16-13).

2. Random Noise and Contrast-to-Noise Ratio

Random noise, also called *statistical noise* or *quantum mottle,* is present everywhere in a nuclear medicine image. Even when the size of an object is substantially larger than the limiting spatial resolution of the image, statistical noise can impair detectability, especially if the object has low contrast. The critical parameter for detectability is the CNR of the object in the image. In the following discussion, we present an analysis and some illustrations of the effects of CNR on detectability of objects in 2-D planar nuclear images.

Suppose that a 2-D image contains a circular lesion of area A_ℓ having contrast C_ℓ (Equation 15-6) against a uniform background counting rate, R_o (cps/cm^2). The number of counts recorded in a background area of the same size as the lesion during an imaging time, t, is

$$N_o = R_o \times A_\ell \times t = R_o \times \frac{\pi}{4} d_\ell^2 \times t \quad (15\text{-}8)$$

where d_ℓ is the diameter of the lesion. The statistical variation of counts in background areas of size A_ℓ is

$$\sigma_{N_o} = \sqrt{N_o} = \sqrt{R_o \times \frac{\pi d_\ell^2}{4} \times t} \quad (15\text{-}9)$$

Thus the fractional standard deviation of counts due to random statistical variations is

$$C_{noise} = \frac{\sigma_{N_o}}{N_o} = \frac{1}{\sqrt{R_o \times \frac{\pi d_\ell^2}{4} \times t}} \quad (15\text{-}10)$$

As indicated by the notation in Equation 15-10, this factor can be considered as the "noise contrast" for a circular area of diameter d_ℓ in background areas of the image. The ratio of lesion-contrast to noise-contrast is defined as its CNR_ℓ.

$$CNR_\ell = \frac{|C_\ell|}{C_{noise}} \approx |C_\ell| \times d_\ell \times \sqrt{R_o \times t} \approx |C_\ell| \times d_\ell \times \sqrt{ID_o} \quad (15\text{-}11)$$

where we have used the approximation $\sqrt{\pi/4} \approx 1$. The quantity $ID_o = (R_o \times t)$ is the background *information density* of the image and has units (counts/cm^2). The absolute value of contrast, $|C_\ell|$, is used in Equation 15-11 to indicate that it applies to either positive or negative contrast.

To detect a lesion or other object in an image, the observer must be able to distinguish between the lesion or object and noise-generated contrast patterns in background areas of the same size in the image. A substantial amount of research has gone into this subject. The conclusion is that, to be detectable, an object's CNR must exceed 3-5. This factor is known as the *Rose criterion,* after the individual who did basic studies on this subject.[4] The actual value depends on object size and shape, edge sharpness, viewing distance, observer experience, and so forth. Choosing a factor of 4, the requirement for detectability becomes $CNR_\ell \geq 4$, and Equation 15-11 can be written as

$$|C_\ell| \times d_\ell \times \sqrt{R_o \times t} \geq 4 \quad |C_\ell| \times d_\ell \times \sqrt{ID_o} \geq 4 \quad (15\text{-}12)$$

Equation 15-12 applies to somewhat idealized conditions of more or less circular objects against a relatively uniform background of nonstructured noise. Such conditions rarely apply in nuclear medicine. Nevertheless, this equation can be used to gain some insights into lesion detectability and the factors that affect it.

EXAMPLE 15-2

Estimate the minimum contrast for detection of circular objects of 1-cm and 2-cm diameter in an area of an image where the background information density is ID_o = 400 counts/cm^2.

Answer

Rearranging Equation 15-12 and inserting the specified information density,

$$|C_\ell| \geq \frac{4}{d_\ell \sqrt{ID_o}} = \frac{4}{\sqrt{400} d_\ell} = \frac{0.2}{d_\ell}$$

Thus for a 1-cm diameter object, the minimum contrast required for detectability is approximately 0.2 (20%), whereas for a 2-cm diameter object it is approximately 0.1 (10%).

Example 15-2 shows that, all other factors being the same, the contrast required for detectability is inversely proportional to object size.

EXAMPLE 15-3

Estimate the minimum diameter for detection of an object that has 10% contrast, $|C_\ell| = 0.1$, in an area of the image where the background information density is 100 counts/cm^2.

Answer

Rearranging Equation 15-12 and inserting the specified parameters,

$$d_\ell \geq \frac{4}{|C_\ell| \sqrt{ID_o}} = \frac{4}{0.1\sqrt{400}} = 4 \text{ cm}$$

Examples 15-2 and 15-3 illustrate that the minimum size requirement for object detectability decreases inversely with the square root of information density, from 2 cm with ID_o = 400 counts/cm^2 in Example 15-2 to 4 cm with ID_o = 100 counts/cm^2 in Example 15-3.

At first glance, it would seem that adding background radiation to an image would improve lesion detectability, by increasing information density, ID_o. However, as illustrated by Example 15-1, background radiation also degrades lesion contrast. The following

example illustrates the overall effect of background radiation on object detectability.

EXAMPLE 15-4

Example 15-3 indicates that a 4-cm diameter object with 10% contrast should be detectable against a background information density of ID_o = 100 counts/cm^2. Suppose that background radiation with the same information density, (ID_b = 100 counts/cm^2) is added to the image. Estimate the minimum detectable lesion size after this is done.

Answer

According to Example 15-1, the addition of background radiation with $ID_b = ID_o$ decreases contrast by a factor of 2, from $C_\ell = 0.1$ to $C'_\ell = 0.05$. At the same time, the background information density increases from 100 counts/cm^2 to 200 counts/cm^2. Rearranging Equation 15-12 and inserting these values, one obtains

$$d_\ell \geq \frac{4}{|C'_\ell|\sqrt{ID_o}} = \frac{4}{0.05\sqrt{200}} \approx 5.7 \text{ cm}$$

Example 15-4 illustrates that the minimum detectable object size becomes larger (from 4 cm to 5.7 cm in the example) when background radiation is added. This example illustrates that the effect of background on degradation of object contrast more than offsets its effect toward increasing the information density of the image.

These examples and analyses assume that object size (d_ℓ) and contrast (C_ℓ) are independent variables. This may be true for computer-generated test images; however, in planar nuclear medicine imaging these parameters often are intimately linked. In many cases, lesions are somewhat spherical in shape so that their thickness varies linearly with their diameter. Thus a larger lesion not only has a larger diameter but generates greater contrast as well. The following example illustrates how these two factors operating together affect lesion CNR and detectability.

EXAMPLE 15-5

Suppose that a certain radionuclide concentrates in normal tissue to a level that provides a counting rate of 10 cpm/cm^2 per cm of tissue thickness. Suppose further that it concentrates in a certain type of lesion to a level that is twice this value, that is, 20 cpm/cm^2 per centimeter of lesion thickness. Compare the contrast, CNR and detectability of 1-cm diameter versus 2-cm diameter lesions embedded in normal tissue of total thickness 10 cm and an imaging time of 1 min. Ignore the effects of attenuation and source-to-detector distance for this comparison.

Answer

In both cases, the uptake in normal tissues would generate a background counting rate of $R_o = 10 \text{ cm} \times 10 \text{ cpm/cm}^2$ per centimeter thickness = 100 cpm/cm^2. For the 1-cm diameter lesion, the count rate over the center of the lesion is

$$\begin{aligned}&(9 \text{ cm} \times 10 \text{ cpm/cm}^2 \text{ per cm})\\ &\quad + (1 \text{ cm} \times 20 \text{ cpm/cm}^2 \text{ per cm})\\ &\quad = 110 \text{ cpm/cm}^2\end{aligned}$$

Thus the contrast of the 1-cm diameter lesion is (110 – 100)/100 = 0.1 (10%). For a 1-min imaging time, its CNR (Equation 15-11) is

$$CNR_{1\,cm} = 0.1 \times 1 \times \sqrt{100 \times 1} = 1$$

For the 2-cm diameter lesion, the counting rate over the center of the lesion is

$$\begin{aligned}&(8 \text{ cm} \times 10 \text{ cpm/cm}^2 \text{ per cm})\\ &\quad + (2 \text{ cm} \times 20 \text{ cpm/cm}^2 \text{ per cm})\\ &\quad = 120 \text{ cpm/cm}^2\end{aligned}$$

Thus the contrast of the 2-cm diameter lesion is (120 – 100)/100 = 0.2 (20%). For a 1-min imaging time, its CNR is

$$CNR_{2\,cm} = 0.2 \times 2 \times \sqrt{100 \times 1} = 4$$

According to Equation 15-11, when the diameter of a *planar* object is doubled, its CNR increases by a factor of 2 as well. However, Example 15-5 shows that when the object is spherical, so that its thickness is doubled as well, its CNR increases by another factor of 2, that is, the total change in CNR is a factor of 4. This example illustrates the strong dependence of lesion detectability on lesion size when its contrast increases with its size. In essence, CNR increases as the *square* of spherical lesion diameter, not as the first power as implied by Equation 15-11. Going in the opposite direction, this factor becomes a significant impediment for the detection of smaller and smaller lesions in nuclear medicine.

Finally, Equations 15-11 and 15-12 and the discussion thus far assume that detectability is the same for positive ("hot spot") and negative ("cold spot") contrast. Indeed, if two objects generate identical levels of contrast, this is a valid assumption. Again, however, additional factors come into play in nuclear

medicine imaging. Specifically, the intrinsic contrast of a lesion can depend on whether its contrast is generated by preferential uptake or by preferential suppression of uptake relative to surrounding normal tissues. The following example provides an illustration.

EXAMPLE 15-6

Suppose that two radiopharmaceuticals are available for a study. Radiopharmaceutical A generates contrast by selective uptake in a lesion that is 10 times higher than the uptake in surrounding normal tissue, whereas radiopharmaceutical B generates contrast by suppression of uptake in the same lesion, to a level that is 1/10 (10%) of the uptake in surrounding tissue. Thus the uptake ratio is 10:1 in both cases. Assume that a 1-cm thick lesion is present in a total thickness of 10 cm of tissue. Ignoring the effects of attenuation and source-to-detector distance, calculate the CNR generated by the two radiopharmaceuticals. For both radiopharmaceuticals, assume that the uptake in normal tissue generates a counting rate of 10 cpm/cm^2 per centimeter thickness of tissue and that imaging time is 1 min in both cases.

Answer

For both radiopharmaceuticals, the uptake in normal tissues generates a background counting rate of R_o = 10 cm × 10 cpm/cm^2 per centimeter thickness = 100 cpm/cm^2. For radiopharmaceutical A ("hot" lesion), the uptake of the lesion is 10 times greater and the counting rate over the lesion is

$$(9\text{ cm} \times 10\text{ cpm/cm}^2\text{ per cm}) + (1\text{ cm} \times 100\text{ cpm/cm}^2\text{ per cm}) = 190\text{ cpm/cm}^2$$

Thus the contrast of the hot lesion is (190 – 100)/100 = 0.9 (90%). For a 1-min imaging time, its CNR (Equation 15-11) is

$$\text{CNR}_\text{A} = 0.9 \times 1 \times \sqrt{100 \times 1} = 9$$

This value easily exceeds the requirement for detectability given by the Rose criterion.

For the "cold" lesion, the uptake by the lesion is 1/10 of the uptake in normal tissue. Thus the counting rate over the lesion is

$$(9\text{ cm} \times 10\text{ cpm/cm}^2\text{ per cm}) + (1\text{ cm} \times 1\text{ cpm/cm}^2\text{ per cm}) = 91\text{ cpm/cm}^2$$

Thus the contrast of the "cold" lesion is (91 – 100)/100 = –0.09 (–9%), in which the minus sign indicates "negative" contrast. For a 1-min imaging time, the CNR for radiopharmaceutical B is

$$\text{CNR}_\text{B} = 0.09 \times 1 \times \sqrt{100 \times 1} = 0.9$$

which is well below the threshold of detectability specified by the Rose criterion.

Example 15-6 illustrates the basis for the generally held (and generally accurate) belief that "cold" lesions are more difficult to detect than "hot" ones. One way to overcome this deficit is to inject more radioactivity for "cold" lesions; however, the specified levels of uptake in normal tissue in this example leads to comparable radiation doses in both cases and thus the higher level of radioactivity required for radiopharmaceutical B would presumably lead to greater radiation dose. Specific comparisons of radiopharmaceuticals vary, depending on details of the uptake distribution and properties of the radionuclides involved.

These examples illustrate that contrast and information density can be limiting factors for lesion detection, even when the size of the lesion easily exceeds the spatial resolution limits of the imaging system. Figure 15-12 further illustrates this point for images of a heart phantom. Although spatial resolution and contrast are the same for all the images shown in this figure, there are marked differences in lesion visibility because of differences in information density and noise.

Although not specifically included in the analysis of CNR presented earlier, spatial resolution of the imaging system also affects the detectability of small, low-contrast objects. As shown in Figure 15-13, high-resolution collimators (or imaging detectors) provide better image contrast and improved visibility for fine details, even for smaller numbers of counts in the image. In essence, "sharpening" the edges of lesions lowers the CNR required for detectability (Rose criterion) specified in Equation 15-12.

Nevertheless, the tradeoff between improved collimator resolution and decreased collimator sensitivity (see Equations 14-7 and 14-8), as well as the requirement for greater information density, eventually establishes a point of diminishing returns in the effort to detect smaller and smaller lesions by improvements in imaging resolution. In the end, detectability in nuclear medicine is limited by

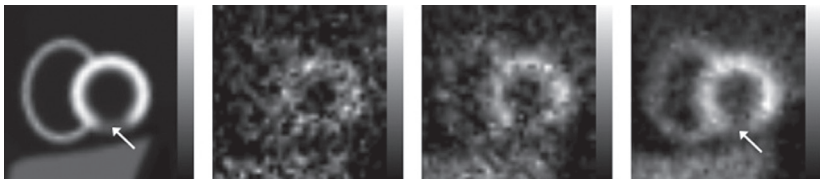

FIGURE 15-12 Example of effects of information density on visibility of a low-contrast lesion (*arrow*) in a computer-simulated cardiac phantom. The simulation assumes a ^{99m}Tc radiotracer imaged on a gamma camera with a 15% energy window. Note the region of reduced radiotracer uptake in the myocardium (*arrow*) that can be clearly visualized only with the highest information density. (*Courtesy Dr. Hendrik Pretorius and Dr. Michael King, University of Massachusetts Medical School, Worcester, MA*).

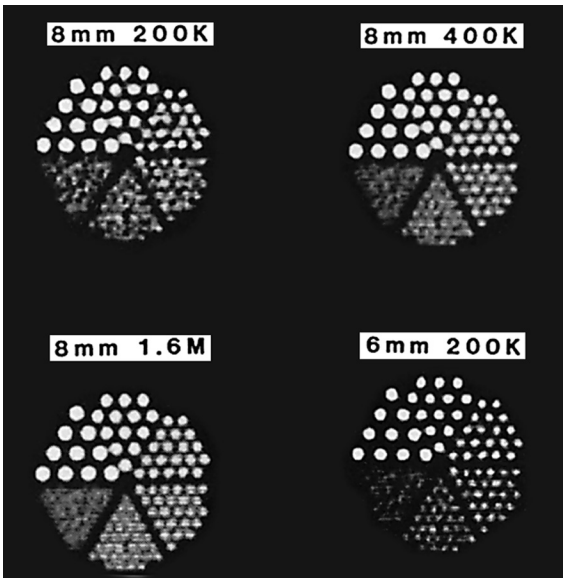

FIGURE 15-13 Demonstration of effects of improved resolution on contrast and detectability of small objects. Improved spatial resolution results in improved contrast (*lower right*), providing improved visibility in spite of fewer counts in comparison with the other images. Decreased sensitivity of high-resolution collimators ultimately sets practical limits for high-resolution imaging in nuclear medicine. (*From Muehllehner G: Effect of resolution improvement on required count density in ECT imaging: A computer simulation.* Phys Med Biol *30:163-173, 1985.*)

information density rather than image resolution. Information densities in planar nuclear medicine images are typically in the range of 100 to 3000 counts/cm^2. This is well below the levels encountered in radiography and photography, in which information densities (x-ray or visible light photons detected to form the image) are on the order of 10^6 events per mm^2. Practical limitations on imaging time and the amount of activity that can be administered safely to patients are serious impediments to improvements in nuclear medicine information densities and are the reason why photographic or radiographic image quality are unlikely to ever be achieved in nuclear medicine.

In general, the rules regarding image CNR and object detectability are the same for planar images and tomographic images; however, the approaches for calculating CNR are different. This is discussed further in Chapter 16, Section C.3.

E. OBSERVER PERFORMANCE STUDIES

The physical measures of image quality discussed in the preceding sections are helpful for comparing different imaging systems, as well as for preparing purchase specifications, establishing quality assurance parameters, and so forth. They also can in some cases provide useful estimates of minimum detectable object size and contrast, as in Examples 15-2 through 15-6. In most cases, however, object detectability is determined more accurately by direct evaluation, using human observers. The general name for such evaluations is *observer performance studies.* They test both the ability of an imaging device to produce detectable objects as well as the ability of individual observers to detect them. Two types of experiments commonly used for this purpose in nuclear medicine imaging are *contrast-detail* (C-D) and *receiver operating characteristic* (ROC) studies.

1. Contrast-Detail Studies

A contrast-detail, or C-D study is performed using images of a phantom having a set of objects of different sizes and contrasts. Typically, the objects are graded in size along one

axis of the display and in contrast along the other. An example is the Rollo phantom, shown in Figure 15-14A. This phantom consists of solid spheres of four different diameters immersed in four different thicknesses of a radioactive solution of uniform concentration. Images of this phantom thus contain cold lesions of different sizes and contrasts (Fig. 15-14B).

To perform a C-D study with this or a similar phantom, images are obtained using the different imaging systems or techniques to be evaluated. An observer then is given the images, usually without identification and in random order to avoid possible bias, and asked to indicate the smallest diameter of sphere that is visible at each level of contrast. Borderline visibility may be indicated by selecting a diameter between two of the diameters actually present in the image. The results then are presented on a C-D diagram as illustrated by Figure 15-15. A C-D study can be helpful for comparing detectability of both large low-contrast lesions as well as small high-contrast lesions. For example, in Figure 15-15, system A would be preferred for the former and system B for the latter. Because of the subjective nature of C-D studies, the use of multiple observers is recommended. Also, because observers may change their detection threshold from one study to the next or as they gain familiarity with the images, it usually is helpful to repeat the readings for verification of results.

C-D studies have a number of limitations. Because they are subjective, they are susceptible to bias and other sources of differences in the observer's detection thresholds in different experiments. This is especially true for phantoms having a design similar to the one illustrated in Figure 15-14, because the observer has a priori knowledge of the locations of the simulated lesions. Thus such a phantom does not test for the possibility of false-positive results, that is, the mistaken detection of objects that actually are not present in the image. This is particularly important for noisy images in which noise not only can mask the presence of real objects but also can create apparent structures that masquerade as real objects. Finally, C-D phantoms generally are lacking in clinical realism.

2. Receiver Operating Characteristic Studies

Some of the deficiencies of the C-D method outlined earlier are overcome by the ROC method. For an ROC study, a set of images is obtained with the different imaging systems or techniques to be tested. Phantoms containing simulated lesions can be used, but it also is possible to use actual clinical images. In the simplest approach, each image contains either one or no lesions. The former are called *positive images* and the latter are called *negative images.* The images are given to the observer, who is asked to indicate whether a lesion is present or absent in each image, as well as where it is and his or her confidence that it actually is present. Usually the confidence levels are numbered and four different levels are permitted; for example, 1 = definitely present,

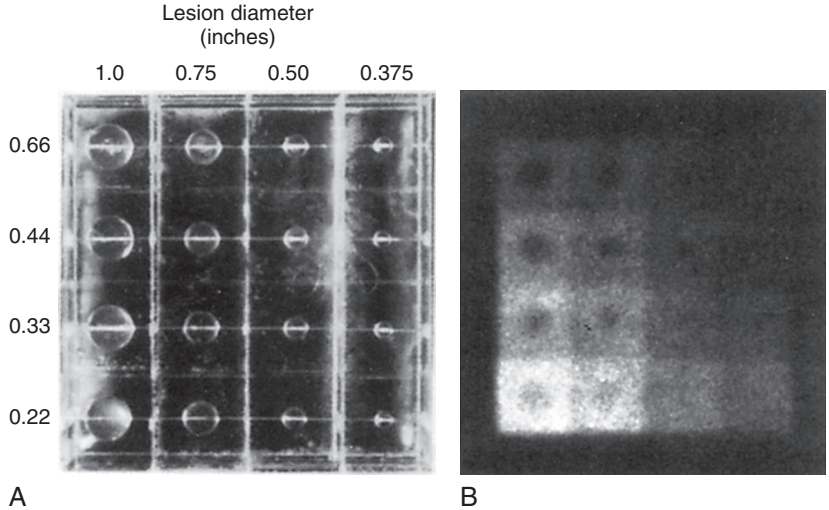

FIGURE 15-14 Example of a phantom, the Rollo phantom, which can be used to obtain images for a contrast-detail study. *A,* Phantom. *B,* Example image. (*From Rollo FD, Harris CC: Factors affecting image formation. In Rollo FD [ed]:* Nuclear Medicine Physics, Instrumentation, and Agents. *St. Louis, 1977, CV Mosby, p 397.*)

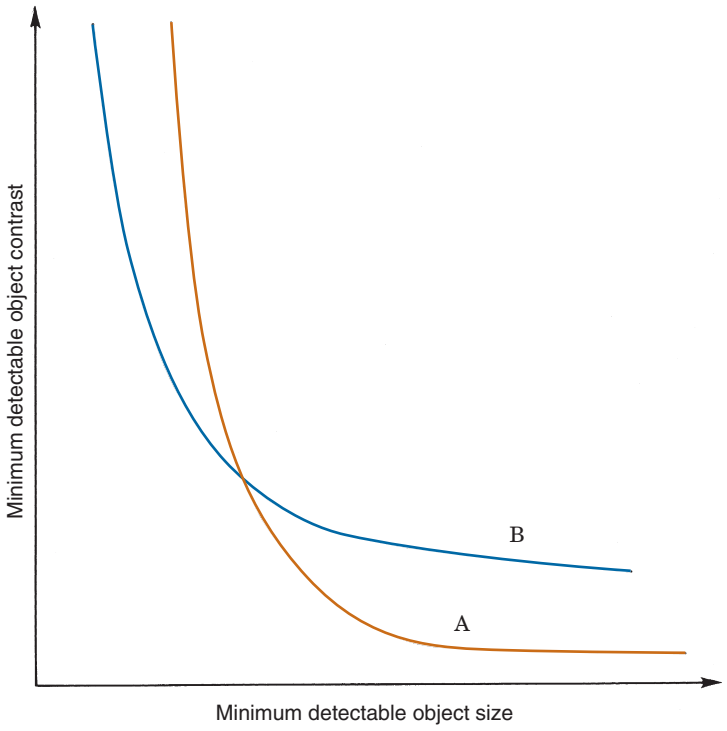

FIGURE 15-15 Hypothetical results of a contrast-detail study comparing two imaging systems, A and B. Reading horizontally, one can estimate the minimum size of an object that can be detected for a specified level of object contrast. Reading vertically, one can estimate the minimum contrast required for detection of an object of a specified size. In this example, system A provides better detectability for large low-contrast objects, suggesting perhaps a better lesion-to-noise contrast ratio, whereas system B is better for small high-contrast lesions, suggesting perhaps better spatial resolution.

2 = probably present, 3 = probably not present, and 4 = definitely not present. Then the following results are calculated for each confidence level:

True-positive fraction (TPF) = fraction of positive images correctly identified as positive by the observer

False-positive fraction (FPF) = fraction of negative images incorrectly identified as positive by the observer

Two other parameters that are calculated are the *true-negative fraction* (TNF) = (1 – FPF), and the *false-negative fraction* (FNF) = (1 – TPF). The TPF is sometimes called the *sensitivity* and TNF the *specificity* of the test or the observer.

The ROC curve then is generated by plotting TPF versus FPF for progressively relaxed degrees of confidence, that is, highest confidence = level 1 only, then confidence levels 1 + 2, then confidence levels 1 + 2 + 3, and so forth. An example of data and the resulting ROC curves are shown in Figure 15-16. The ROC curve should lie above the ascending 45-degree diagonal, which would represent "guessing." The farther the curve lies above the 45-degree line, the better the performance of the imaging system and observer.

An ROC curve shows not only the true-positive detection rate for an observer or an imaging system or technique but also its relationship to the false-positive detection rate. Thus it is relatively immune to the sources of observer bias that can occur in C-D studies, for example, a tendency to "over-read" to avoid missing a possible lesion or test object. It also is applicable to other types of detection questions, such as the presence or absence of disease, which might be indicated by a general pattern of uptake within an organ, as opposed to the simple detection of individual lesions.

As with C-D studies, the interpretation of ROC results sometimes can be challenging. For example, the ROC curves for two different imaging systems can "cross," leading to some ambiguity in the results. One approach to simplifying the interpretation is to report the

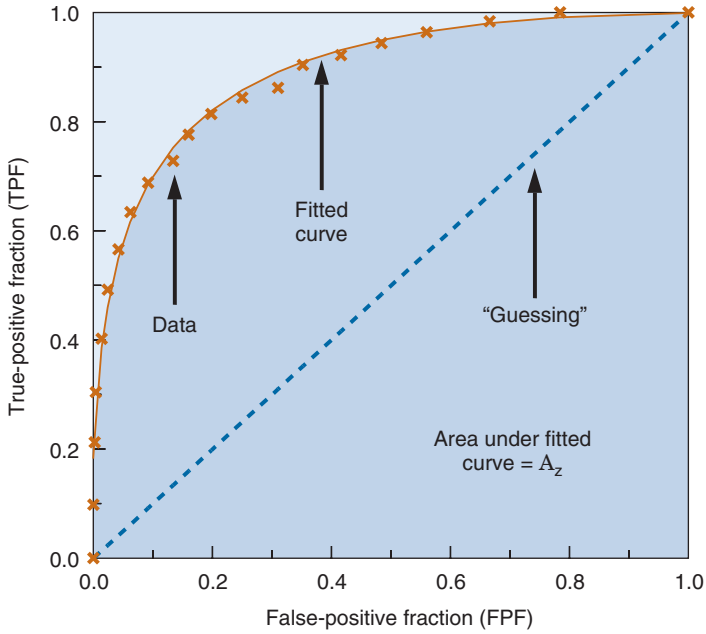

FIGURE 15-16 Example of results from an ROC study. ✖, data points; *orange line*, fitted curve; *blue line,* 45-degree line, which is equivalent to "guessing." Area under the curve (*shaded in darker blue*) is A_z, which is one measure of detection accuracy.

results of an ROC study as a single number. Most commonly the parameter calculated is the area under the ROC curve, usually denoted by A_z. This number can range from zero (all readings wrong) to 1 (all readings correct). A value of 0.5 indicates an overall accuracy of 50%, which is equivalent to "guessing."

An extensive amount of theoretical and experimental work has been done on the properties of A_z, including such issues as statistical comparisons of values obtained from different ROC studies. References 5 to 7 present detailed analyses of these and other practical issues in ROC studies. A_z also has an interesting practical interpretation: It is the probability that, given a side-by-side pair of images, one of which has a lesion or test object and the other does not, the observer will correctly identify the image with the lesion.[8]

Despite their power and potential usefulness, ROC studies also have a number of limitations. Perhaps the most challenging is the verification of absolute "truth" for images obtained from clinical studies. Ideally, the outcome of the ROC study itself (i.e., the tested images) should not be used for this determination. This means that other equally or even more reliable information about the presence or absence of disease in the patient must be available. Often nonimaging tests (e.g., surgical results) must be obtained for verification of "truth" when clinical images are used.

Another potential problem is the possible presence of multiple lesions on a single image. Conventional ROC methodology allows only for a single "yes-no" interpretation of each image. This allows straightforward calculations of false-positive rates. However, if multiple lesions are possible, as in many clinical images, the potential number of false positives is virtually infinite, making the calculation of false-positive rates difficult, if not impossible. Alternative methods, called the *free-response operating characteristic*, that allow for the presence of multiple lesions have been developed and are discussed in reference 9.

Finally, even a "perfect" image evaluation technique with a clearly defined outcome might not provide the final answer regarding the merit or value of an imaging device or technique. Even after the physician or scientist has demonstrated that he or she has developed a truly "better" device or technique in terms of lesion or disease detectability, there is still the bottom-line question: "So what?" Does the improved detectability alter the care of the patient or the outcome of that care? Does it improve the patient's quality of life? In an age of cost-consciousness, what are the cost-benefit tradeoffs? For example, from a public health perspective, is it really worth

spending a small fortune to detect the next smaller size of lesion, as compared with directing those funds toward simpler health measures, such as education and behavior modification? These are difficult questions to answer, but efforts are being made to develop methodology for answering them in a quantitative and objective manner. The general term for these investigations is *efficacy studies*. Additional discussion of this topic can be found in reference 10.

REFERENCES

1. Barrett HH, Yao J, Rolland JP, Myers KJ: Model observers for assessment of image quality. *Proc Natl Acad Sci* 90:9758-9765, 1993.
2. Cunningham IA: Introduction to linear systems theory. In Beutel J, Kundel HL, Nan Metter RL, editors: *Handbook of Medical Imaging*, Bellingham, WA, 2000, SPIE, Chapter 2.
3. Vayrynen T, Pitkanen U, Kiviniitty K: Methods for measuring the modulation transfer function of gamma camera systems. *Eur J Nucl Med* 5:19-22, 1980.
4. Rose A: *Vision: Human and Electronic*. New York, 1973, Plenum Press, pp 21-23.
5. Swets JA, Pickett RM: *Evaluation of Diagnostic Systems: Methods from Signal Detection Theory*, New York, 1982, Academic Press.
6. Metz CE: ROC methodology in radiologic imaging. *Invest Radiol* 21:720-733, 1986.
7. Metz CE: Fundamental ROC analysis. In Beutel J, Kundel HL, Nan Metter RL, editors: *Handbook of Medical Imaging*, Bellingham, WA, 2000, SPIE, Chapter 15.
8. Hanley JA, McNeil BJ: The meaning and use of the area under the receiver operating characteristic (ROC) curve. *Radiology* 143:29-36, 1982.
9. Chakraborty DP: The FROC, AFROC, and DROC variants of the ROC analysis. In Beutel J, Kundel HL, Nan Metter RL, editors: *Handbook of Medical Imaging*, Bellingham, WA, 2000, SPIE, Chapter 16.
10. Fryback DG, Thornbury JR: The efficacy of diagnostic studies. *Med Decis Making* 11:88-94, 1991.