Specification of Performance of Positron Emission Tomography Scanners

In the process of selecting a positron emission tomography scanner, we were hampered by the lack of standardized methods for measuring instrument performance. The assessment of resolution and sensitivity illustrates the problem.

Resolution is usually stated as the full width at half maximum (FWHM) of the count profile (line-spread function) through the reconstructed image of a line source (*1*). Several variables (often unspecified) can affect the FWHM. First, as the diameter of the experimental "line" source increases, the measured FWHM increases, since the observed line-spread function is the convolution of the true line-spread function with the rectangular function describing the width of the source. Second, the measured FWHM is less if the line source is imaged in air rather than in a scattering medium, both because of the lack of scattered radiation in air (which mainly increases the "tails" of the line-spread function) and because the positron range is greater. In air, the positrons are annihiliated either in the line source itself or at a great distance from it. In a scattering medium, annihilation events occur in and adjacent to the line source (*2,3*). The measured FWHM is less with the line source in air, since the observed line-spread function is the convolution of the true line-spread function with the positron range distribution function, which is much broader in a scattering medium. The use of a metal needle rather than plastic tubing for the line source also decreases the average positron range, decreasing the measured FWHM. The measured FWHM also depends on the radionuclide, which has a characteristic maximum positron energy. Lower-energy positrons have a smaller range in scattering material (*4*). The reconstruction algorithm itself affects the measured FWHM (*5*). For example, a ramp filter will provide maximum resolution, but will generate artifacts such as overshoot and ringing ("Gibbs phenomena") at borders of objects. Filters that do not produce artifacts yield a larger measured FWHM.

As a step in the direction of standardization, we recommend the use of plastic tubing with an inner diameter of 1 mm or less. The tubing should be filled with Ga-68 in water, since this radionuclide has an intermediate positron energy and is readily available, and placed in a plastic phantom with 20 cm o.d. filled with water. Since measured resolution is limited by the range of the higher-energy positron of Ga-68, a second measurement should also be done with F-18, which has a lower-energy positron. Measurements should be made with the line source at various positions that encompass the entire field of view. A Shepp and Logan filter should be used in the reconstruction, since it yields high resolution without artifacts. Both the full width at half maximum and full width at tenth maximum (as an indication of scatter) should be stated.

Sensitivity is usually expressed as the observed counts per second per microcurie per milliliter from a 20-cm-diameter phantom filled with water (*6*). As with resolution, several variables can affect this measurement, and it should be standardized. First, the phantom diameter should be specified as being the internal or external diameter, since this will affect the total amount of water and therefore the total amount of radioactivity in the phantom. Second, the wall thickness of the phantom should be standardized. The thicker the wall the greater the attenuation and the lower the observed count rate. Third, although the count rate is expressed "per microcurie per ml," measurements are usually made with lower concentrations of radioactivity. A lower concentration with a correspondingly lower count rate will yield more observed counts per second when normalized to 1 μCi/ml because dead-time problems will not be as great. The axial length of the phantom should also be standardized. The longer the phantom the more out-of-plane scatter radiation will enter the slice and increase the observed counts per second. Some describe sensitivity as the "raw" observed count rate; others normalize for differing detector efficiencies, and make scatter, coincidence, or attenuation corrections. Scatter and random coincidence corrections decrease the final count rate, but normalization and attenuation corrections increase it. Slice thickness (axial field of view per slice) also affects the measured sensitivity. As the slice thickness increases, the volume of water and therefore the amount of radioactivity increases, increasing the observed count rate without actually changing the "efficiency" of detection.

We recommend the use of a phantom with 1/8″ thick walls, filled with Ga-68 in water at a concentration of 100 nCi/ml. We prefer a phantom 15 cm in diameter for head scanners, and 25 cm in diameter for whole-body scanners. The phantom should be 2 cm longer than the axial field of view of the slice. The count rate should be the "raw" observed total coincidence rate, before normalization or other corrections. The count rate should be expressed on a "per axial cm" basis, using the measured axial FWHM of a line source in the plane of the slice to normalize the observed count rate.

Even if well standardized, resolution and sensitivity measurements are not sufficient to fully characterize the performance of a positron emission tomograph. Other specifications, such as accidental and random coincidence count rates as a function of total coincidence count rates, axial resolution, and quantitative linearity, are also needed. While we have not attempted to suggest standards for the specification of all parameters, standardization of certain fundamental measurements of performance could be of value.

JONATHAN M. LINKS
HENRY N. WAGNER, JR.
The Johns Hopkins Medical Institutions
Baltimore, Maryland

REFERENCES

1. HOFFMAN EJ, PHELPS ME, MULLANI NA, et al: Design and performance characteristics of a whole-body positron transaxial tomograph. *J Nucl Med* 17:493–502, 1976

2. PHELPS ME, HOFFMAN EJ, HUANG SC, et al. Effect of positron range on spatial resolution. *J Nucl Med* 16:649–652, 1975

3. CHO ZH, CHAN JK, ERICKSSON L, et al: Positron ranges obtained from biomedically important positron-emitting radionuclides. *J Nucl Med* 16:1174–1176, 1975

4. GRAHAM LS, MACDONALD NS, ROBINSON GD, LLACER J: Effect of positron energy on spatial resolution. *J Nucl Med* 14:401–402, 1973 (abst)

5. BUDINGER TF, GULLBERG GT: Transverse section reconstruction of gamma-ray emitting radionuclides in patients. In *Reconstruction Tomography in Diagnostic Radiology and Nuclear Medicine*. M. M. Ter-Pogossian, Ed. Baltimore, University Park Press, 1977, pp 315–342

6. PHELPS ME, HOFFMAN EJ, HUANG SC, et al: ECAT: A new computerized tomographic imaging system for positron-emitting radiopharmaceuticals. *J Nucl Med* 19:635–647, 1978

Comments on Specifying the Performance of a Positron Tomograph

The authors of the preceding letter have presented an excellent list of caveats for the potential user/buyer of a positron tomograph, and their suggestions for standardization are excellent as far as they go. It should be remembered, however, that the tomograph user/buyer must invest a great deal in the tomograph: either one to two years of his time if he builds it, or 0.5 to 2.0 million dollars if he buys it. Unless the user is primarily a builder/designer of instruments, he receives little or no scientific credit for building or evaluating the device. His greatest return will come from his application of the tomograph to medical investigations. The performance of the device should be comprehensively and critically evaluated to protect his investment. Dependence upon inadequately defined specifications of resolution and sensitivity can lead one to purchase or build a tomograph that is incapable of performing the measurements for which it was acquired.

Resolution. The intrinsic resolution of a pair of coincidence detectors is best midway between the two detectors, and deteriorates with distance from the midpoint. The smaller the separation between the detectors, the faster the resolution deteriorates with distance from the midpoint, but the resolution at the midpoint is identical for most separations. A specification differentiating between the resolutions of two systems of different diameters but identical detectors requires a knowledge of the resolution as a function of distance from the center of the field of view (FOV). An example of the magnitude of this effect is the variation in resolution of a pair of coincident detectors 17 mm wide by 28 mm high: it varies from 9.8 to 13.0 mm FWHM (33%) from midpoint to a 12-cm radius when the detector separation is 45 cm (*1*) (this is a system diameter comparable to some commercially available tomographs). Geometrically, the fraction of the area of the FOV viewed by the midpoint of coincident detector pairs is small compared with the area viewed at a distance from the midpoint. This means that the lower-resolution region of the detector response, near the edge of the image, dominates the resolution in the image more than the high-resolution response seen at the center. From these considerations it can be seen that specifying the resolution only at the center can be very misleading. A resolution specification should be very specific about resolution as a function of position in the FOV. At this time in the development of positron tomography, the minimum acceptable specification of resolution should be no less than average resolution over the inner three fourths of the FOV and the variation of the resolution stated as the standard deviation of the mean of the measured values.

It should also be noted that the use of the standard Shepp filter is adequate for comparisons of tomographs, but many clinics use different filters with lower frequency responses in actual imaging procedures. To avoid disappointment with the tomograph's performance in the clinic, the buyer should ascertain the resolution under practical clinical imaging conditions.

Slice thickness. This should also be specified as a function of position in the FOV. This is particularly critical in evaluating multislice systems designed for brain imaging. In these systems, coincidence data are collected between image planes as an approximation of the data for the plane midway between the two normal planes. In simulations of these geometries in our laboratory, assuming a 30-cm patient opening and a 24-cm field of view, we have measured slice thicknesses that vary by a factor of 3 from the center to the edge of the FOV for a variety of system diameters. In this type of image, the data at the edge of the image originate in the two adjacent planes whereas the data at the center are from the correct plane. The validity of such data is obviously compromised. Serious variations in the interplane slice thickness limit the ability to do quantitative measurements, and even simple plane-to-plane comparisons of interplane and regular-plane images are of doubtful validity. These problems are usually suppressed in system evaluations because most test phantoms are thicker than the slice thickness of the tomograph, and consequently are insensitive to variations in that parameter. Single-slice systems may also have significant variations in slice thickness. For instance, the slice thickness in the example cited above varies from 15.7 to 22.5 mm (43%) from center to edge of FOV.

Sensitivity. In specifying sensitivity it is imperative that the accidental-coincidence and scatter-coincidence fractions be measured and specified. The use of the "raw" total coincidence counts can be doubly misleading. First, it includes the accidental and scatter counts as good data and therefore inflates the sensitivity value. Secondly, each accidental or scatter count effectively cancels out two true events in terms of the statistical accuracy of an image (*2*). The error on each scan data point is given by the square root of the number of counts (including scatter and accidentals). Then, when one measures or estimates the scatter and accidentals to allow subtraction of these backgrounds from the total coincidences, there is additional error due to the propagation of the errors in the estimation of the scatter and accidentals. For instance, if the sensitivity measurements are made with a total of 20% scatter and accidentals per average data point (assuming only random counting errors in the determination of scatter and accidentals), then the statistical accuracy is equivalent to the accuracy of a tomograph with 28% lower sensitivity, but with lower background of only 10% accidentals and scatters.

In multislice systems, the sensitivity of the interplane slice must be considered in light of the uniformity of the slice thickness. Since the interplane image is measured with two detector planes, it usually comprises more than 50% of the total system sensitivity. However, if there is a great deal of distortion and inconsistency in these data, the interplane data add little that is worthwhile to the sensitivity of the tomograph.

Total system performance. While good resolution and sensitivity are important, it is also important that all the parts work together to give a good-quality, artifact-free image. Two test objects can provide quick but highly sensitive tests of total system performance. First, a uniform cylinder scanned with very high total counts tests the system for problems with normalization (uniform field correction) or problems in the handling of scatter or accidental events in either software or hardware. These problems will tend to cause circular artifacts in the image. The high photon flux is necessary to ensure that the problem is due to the instrument rather than to poor statistics. These errors are due to discontinuities in the scan data (*2*). Secondly, scanning a complex phantom, such as the pie phantom described by Derenzo (*3*), can allow an evaluation of the resolution uniformity and adequacy of the spatial sampling of the tomograph. The phantom consists of a cylinder of six triangular areas of closely packed holes filled with the positron emitter, each set of different size and spacing. Ideally the image should show the holes in each set as round, equally sized and spaced hot spots. Elongated or different-sized spots indicate nonuniform resolution. Variable intensity or asymmetry in the spots indicates a problem in the spatial sampling. Images of this phantom provide an excellent visual demonstration of both the resolution and overall image quality of a tomograph.

Considering the investment involved in acquiring a positron tomograph, it is worthwhile to consider the specifications in a careful and thorough manner, from resolution to sensitivity to whether or not all the parts actually work together to provide an acceptable image or measurement.

EDWARD J. HOFFMAN
University of California
Los Angeles, California

REFERENCES

1. RICCI AR, HOFFMAN EJ, PHELPS ME, et al: Emission computed tomography simulator. *IEEE Trans Nucl Sci* NS-27:1368–1374, 1980

2. HOFFMAN EJ, HUANG SC, PHELPS ME, et al: Quantitation in positron emission computed tomography. 4. Effect of accidental coincidences. *J. Comput Assist Tomogr* 5:391–400, 1981

3. DERENZO SE, BUDINGER TF, CAHOON JL, et al: The Donner 280-crystal high resolution positron tomograph. *IEEE Trans Nucl Sci* NS-26:2790–2793, 1979