INVITED COMMENTARY

Interpreting Results from a Comparative Study of Lesion Detectability for 6 Different PET Systems

The article by Kadrmas and Christian (*1*), which appears in this issue of *The Journal of Nuclear Medicine*, is a well-written account of a comprehensive study that compares lesion detectability of 6 different commercial PET imaging systems, including 2 state-of-the-art high-resolution dedicated bismuth germanate (BGO) systems, an older dedicated BGO system, a dedicated NaI(Tl) system, and 3 NaI(Tl) hybrid systems. The important contributions of the study include a realistic whole-body phantom with simulated lesions of different sizes and contrasts, a well-designed experimental protocol, and the use of a human observer performance study and the localization receiver operating characteristic (LROC) study paradigm in the evaluation of lesion detectability.

The unique whole-body phantom is comprised of a 3-dimensional (3D) Hoffman brain phantom (Data Spectrum Corp., Hillsborough, NC), an anthropomorphic thorax phantom (Radiology Support Devices Inc., Long Beach, CA), and an elliptic cylinder pelvis phantom (Data Spectrum). The organ compartments are filled with relative activity concentrations to simulate ^{18}F-FDG distribution in patients. Twenty-seven spheric lesions filled with ^{22}Na activity and with inner diameters of 7, 8, 12, and 16 mm are placed at various locations throughout the whole-body phantom. The lesion-to-background activity concentration ratios are 4, 6, 10, and 16 to simulate different lesion contrasts. The elaborate and realistic whole-body phantom provides a realistic simulation of a normal-size patient found in tumor PET studies.

Because the PET systems included in this study came from several categories of PET systems with very different costs and performance characteristics, comparison among them is a complex issue. To minimize experimental variations over time, data from all PET systems were acquired over a period of several weeks. The same 2-dimensional (2D) data acquisition mode was used in all PET systems except for the C-PET system, where the 3D mode was used. To conform to clinical practice, manufacturer-supplied image reconstruction software and manufacturer-suggested default processing parameters in data processing were used. Also, similar data acquisition times were adopted for the same patient throughput and for patient comfort. However, as shown in Table 2 (*1*), the different system sensitivities result in very different total counts in the acquired images.

The lesion detectability performance of the PET systems was evaluated using a localization ROC, or LROC, study design with 11 human observers. Different from the simpler ROC study design, the LROC study paradigm offers a closer resemblance to the lesion detection task in clinical studies. The observer data were analyzed using the LROCFIT program by Swensson (*2*) to obtain LROC curves and the probabilities of correct lesion localization in images obtained from the different PET systems by the human observers. The LROC results were analyzed using several statistical tests.

The results of the study were summarized in Figures 4 and 5 (*1*). They clearly indicate the general superior performance of the dedicated PET systems compared with the NaI(Tl)-based hybrid PET systems in terms of lesion detectability. Among the dedicated PET systems, the performance of the state-of-art BGO-based systems is superior to that of the older BGO-based system, which in turn is superior to that of the NaI(Tl)-based system. A point of interest shown in Figure 6 (*1*) is that for the largest (16 mm) lesions, lesion detectabilities among the different PET systems are almost identical. This suggests that for larger lesions, where statistical noise fluctuations are less important, the similar spatial resolution characteristics of the different PET systems rendered similar lesion detectability. Another point of interest is shown in Figure 7 (*1*), where the same iterative ordered-subsets expectation maximization (OSEM) reconstruction algorithm with different iterative numbers, when applied to data obtained from very similar BGO-based dedicated PET systems, can provide reconstructed images with markedly different quality. Specifically, data from the HR+ PET system (CTI PET Systems, Knoxville, TN), when processed with 7 iterations of the OSEM algorithm, yielded image quality that was superior to that obtained using 2 iterations and was comparable with that from the Advance PET system (General Electric Medical Systems, Milwaukee, WI).

An important note, as indicated by Kadrmas and Christian (*1*), is the exercise of caution in the interpretation of the results of this study. The lesion detectability performance of a PET system depends on a variety of factors, including the performance characteristics of the system—for example, system sensitivity, spatial resolution and counting rate capability, 2D versus 3D

Received Jul. 25, 2002; accepted Jul. 25, 2002.
For correspondence or reprints contact: Benjamin M.W. Tsui, PhD, Department of Radiology, Johns Hopkins University, JHOC 4263B, 601 N. Caroline St., Baltimore, MD 21287-0859.
E-mail: tsui@jhmi.edu

data acquisition mode, image reconstruction methods and their associated parameters, and correction methods for attenuation and random and scatter events. Table 2 (*1*) shows the very different performance characteristics of the 6 commercial dedicated and hybrid BGO-based and NaI(Tl)-based PET systems, which belong to several different categories of PET system designs. These categories of PET systems also represent large system price differences. For example, the system sensitivity of the 2 high-end dedicated PET systems is 3–10 times that of the NaI(Tl)-based hybrid PET system. The counting rate capabilities of the high-end dedicated BGO-based PET systems are much higher than those of the NaI(Tl)-based dedicated and hybrid PET systems. The different counting rate capabilities also result in different optimized system operating parameters and specific activity. As a result, in the experimental design that uses equal data acquisition time or patient throughput in clinical applications, images from the higher-end dedicated PET systems have significantly higher total coincidence counts and lower noise levels than those of the other PET systems. Therefore, it is not too surprising that the high-end dedicated PET images with higher system costs give higher lesion detectability than those of other PET systems with lower costs for most lesion sizes.

An interesting question is: Would one expect similar lesion detectability between images obtained from the different PET systems with similar total coincidence counts? The answer would provide a true comparison of the system performance without the effect of counting statistics. However, because of the large differences in system sensitivity among the PET systems, very different data acquisition times would be required to achieve similar total coincidence counts. Such an experiment would be very difficult to perform.

The question that has important practical clinical relevance is: For a lesion of a particular size, what is the additional data acquisition time that is required for a lower-end PET system to reach the higher lesion detectability found in a higher-end PET system? The results in Figure 6 (*1*) suggest that for larger lesions, where counting statistics are no longer the main image-degrading factor, the additional acquisition time becomes smaller. This type of information will allow clinicians to make a decision on the trade-off between higher system cost and longer acquisition time for similar lesion detectability of a particular size. The results of this study suggest that the cheaper and lower-end hybrid PET systems may be a cost-effective alternative to the more-expensive and higher-end dedicated PET systems in detecting larger lesions. For smaller lesions, the additional acquisition time required for the cheaper and lower-end hybrid PET systems to achieve lesion detectability similar to that of the more-expensive and higher-end dedicated PET systems may be too high to be clinically practical.

Because the studies needed to answer these questions are beyond the scope of this study (*1*), the readers should avoid overinterpretation of its results. Although it is fair to compare different PET systems of the same categories with similar costs for the same data acquisition time, comparison of dissimilar PET systems from different categories and system costs at equal acquisition time should be studied carefully. Cost-effectiveness of the different PET systems in terms of system costs and acquisition time for similar lesion detectability at a particular lesion size may be used in a more meaningful evaluation. Such evaluation will provide a useful assessment of the utility of the different PET systems in a broad range of practical clinical applications.

Benjamin M.W. Tsui, PhD
Johns Hopkins University
Baltimore, Maryland

REFERENCES

1. Kadrmas DJ, Christian PE. Comparative evaluation of lesion detectability for 6 PET imaging platforms using a highly reproducible whole-body phantom with ^{22}Na lesions and localization ROC analysis. *J Nucl Med*. 2002;43:1545–1554.
2. Swensson RG. Unified measurement of observer performance in detecting and localizing target objects on images. *Med Phys*. 1996;23:1709–1725.