Multicenter Comparison of Calibration and Cross Calibration of PET Scanners

Lilli Geworski, PhD[1]; Bernd O. Knoop, PhD[2]; Maike de Wit, MD[3]; Velimir Ivančević, MD[1]; Roland Bares, MD[4]; and Dieter L. Munz, MD[1]

[1]Clinic for Nuclear Medicine, University Hospital Charité, Humboldt University of Berlin, Berlin, Germany; [2]Clinic for Nuclear Medicine, Hannover Medical School, Hannover, Germany; [3]Department of Oncology, University Hospital Eppendorf, Hamburg, Germany; and [4]Department of Nuclear Medicine, University Hospital, Tuebingen, Germany

The purpose of this study was to compare the calibration of PET scanners and their cross calibration to peripheral devices in a multicenter study. **Methods:** Twenty-three dedicated PET scanners were investigated, applying standardized protocols. To ensure exact determination of the activity used, dose calibrators were checked using ^{68}Ge standards. **Results:** Nine of 19 and 11 of 20 scanners displayed an error of <5% in 3-dimensional and 2-dimensional acquisition modes, respectively. Four and 5 scanners displayed an error of 10% in 3-dimensional and 2-dimensional modes, respectively. All other scanners yielded errors of 5% to <10%. Because of hardware and software problems, the measurements performed on 1 scanner could not be adequately analyzed. **Conclusion:** An investigation of calibration is mandatory. Especially for quantitative analyses in clinical multicenter trials, identification of potentially miscalibrated scanners is necessary.

Key Words: PET; calibration; multicenter study

J Nucl Med 2002; 43:635–639

Use of PET with properly designed ring systems is commonly regarded to be feasible for quantitative studies. The problems associated with quantitation were reviewed comprehensively by Bailey (*1*).

Before one starts any quantitative evaluation of a scanner, proper instrument performance must be ensured by following the quality control procedures recommended by the manufacturer. In general, all quantitative procedures ultimately require that the scanner be calibrated and that the peripheral devices (i.e., dose calibrator and well counter) be cross calibrated to the scanner, preferentially in terms of absolute activity. Calibration gains even more importance when data collected and analyzed by different scanners and at different institutions are compared, typically when pooling patient data for multicenter studies.

Calibration is the process of establishing the relationship between the measured count rate per volume and the true activity concentration. The basic calibration method is similar for all dedicated PET scanners and has to be performed for each mode of data acquisition. However, procedures differ depending on scanner manufacturer and type. Ideally, calibration consists of measuring a phantom containing a known and homogeneous activity concentration, preferably determined with the on-site dose calibrator. For cross calibration of the well counter, an aliquot of the phantom content has to be withdrawn (*2*). In practice, in most cases a manufactured calibration phantom is used, containing a certified activity of the long-lived positron emitter ^{68}Ge in a solid matrix, thus preventing withdrawal of a sample to check either the dose calibrator or the well counter. The volume of matrix carrying the activity is not certified, leaving uncertainty about activity concentration. In this case, the recommended procedure is to cross calibrate the matrix volume against another cylindric phantom containing a solution of a short-lived positron emitter, such as ^{18}F, of known activity and volume. The activity is determined using the on-site dose calibrator, and the well counter is then checked using a sample from this phantom. This type of calibration procedure clearly depends on the accuracy of the on-site dose calibrator, but if the same dose calibrator is used to determine the amount of activity injected into the patient, any small deviation in accuracy will be canceled out, provided that the deviation is constant and the complete cross calibration procedure has been followed, including nuclide-specific corrections for branching ratio and decay. The procedural accuracy of the method that is actually used depends on the accuracy of the corrections applied, especially those for attenuation and scatter. Thus, a calibration independent of attenuation and scatter is desirable (*3*).

The aim of this study was to test the accuracy of scanner calibration and to determine the reasons for unusual deviations from reference standards for the scanners of institutions participating in a specific clinical multicenter study.

MATERIALS AND METHODS

From the institutions participating in the clinical multicenter study, 23 scanners of 7 types were investigated: 15 ECAT EXACT, 3 ECAT EXACT HR+, 1 ECAT 951, and 1 ECAT ART (CTI, Knoxville, TN/Siemens Medical Systems, Inc., Hoffman

Received Aug. 29, 2001; revision accepted Jan. 16, 2002.

For correspondence or reprints contact: Lilli Geworski, PhD, Klinik für Nuklearmedizin, Universitätsklinikum Charité, Humboldt-Universität zu Berlin, Schumannstrasse 20/21, 10098 Berlin, Germany.

E-mail: lilli.geworski@charite.de

Estates, IL); 1 Advance (General Electric Medical Systems, Milwaukee, WI); and 1 QUEST and 1 PENN PET (UGM Medical Systems, Inc., Philadelphia, PA). Of these, 16 were operable in both 2-dimensional (2D) and 3-dimensional (3D) modes, 4 were operable in 2D mode only, and 3 were operable in 3D mode only. This mix fairly represents the installed base of PET in Germany.

In a first step, the accuracy of the dose calibrators at all sites was examined by measuring ^{68}Ge standards (AEA Technology QSA GmbH, Braunschweig, Germany) with 42 MBq, 4.03 MBq, and 0.383 MBq, respectively, in 2 mL, calibrated to an overall uncertainty of ±5%. The prechecked dose calibrator was used to homogeneously fill a cylindric phantom of well-known volume (20 cm in diameter and 20 cm in length) with an aqueous solution of ^{18}F (0.025 MBq/mL for scanners of 2D mode and 0.006 MBq/mL for scanners of 3D mode, reflecting associated differences in sensitivity). From each solution, 2 aliquots (2 mL) were withdrawn and counted in the well counter. Subsequently, the phantom was scanned for 60 min using either mode of data acquisition. The complete procedure, including filling of the phantom, was repeated twice.

After application of all corrections (e.g., scatter, dead time, random coincidences, attenuation, and detector normalization) as provided by the standard software of each scanner, images were reconstructed by applying filtered backprojection using a ramp filter with Nyquist frequency cutoff. If available, 3D data were reconstructed using a linear reprojection algorithm with subsequent 3D filtered backprojection (*4*). In all cases, attenuation correction was performed using measured cold transmission data of high statistical precision (scan duration, 30 to 60 min, depending on actual source strength).

All data were analyzed by the same person using a standardized procedure. After visual inspection for image artifacts, plane-to-plane sensitivity variations, and nonuniformities, the calibration of the scanner and its cross calibration to the dose calibrator were checked by comparing the image activity concentration in regions of interest (15 cm in diameter and centered on the phantom) with phantom activity concentration, as calculated from the activity measured by the dose calibrator and phantom volume. Finally, the cross calibration of the well counter to the scanner was tested by comparing well counter measurements of the aliquots with the corresponding image activity concentration.

Our aim was not to compare and stage different types of scanners or manufacturers—although there are differences—but to ensure the quantitative accuracy of all participating instruments. Therefore, the results from the different PET scanners and their associated periphery were documented graphically in a random sequence established beforehand.

RESULTS

Dose Calibrator and Absolute Activity

Most instruments had errors < 5%; 3 had errors between 7% and 8%. Therefore, nearly all calibrators were well inside the 10% error range normally assigned to the accuracy of this class of instruments. Larger errors showed up in only 3 of 21 instruments and were caused by a defective instrument, artificial high readings because of contamination, and an instrument without manufacturer-supplied radionuclide factors, thus lacking a calibrated reading for positron emitters.

Visual Analysis

All images were free of artifacts, with the exception of a single scanner showing concentric circular structures. These occurred when the nonchanging normalization software components were lost during the chain of normalization procedures.

Scanner Calibration and Cross Calibration to Dose Calibrator

In all scanners, differences in image activity concentration between both repeated phantom measurements were <3% in both 2D and 3D acquisition modes. For this reason, the results of just 1 representative phantom measurement are shown (Fig. 1). Errors from PET concentration measurements, as compared with dose calibrator data, were <5% for 11 of 20 scanners operated in 2D mode and <5% for 9 of 19 scanners operated in 3D mode. For 5 scanners in 2D mode and 4 scanners in 3D mode, the error was approximately 10%. For these, a new cross calibration was recommended to improve accuracy.

Because of incorrect calibration, scanners 2, 3, 4, 11, and 14 did not initially qualify for quantitative data analysis. After the problems had been identified and fixed, the measurements were repeated. The success of these corrections is also shown in Figure 1 (black bars). The identified problems ranged from those that were technical, such as an inoperative dead-time correction because of lost normalization software components (scanner 3) and a failed decay correction because of faulty internal time-zone settings (scanner 2), to those that related to user handling errors, such as a failure to perform calibration at all (scanner 14), an incomplete ^{18}F cross calibration (scanner 11), and a phantom–software mismatch (scanner 4). Finally, scanner 21 had to be excluded from quantitative analysis because of major hardware and software problems that could not be corrected during the time frame of the study.

Cross Calibration to Well Counter

The errors in activity concentrations derived from PET images with respect to well counter measurements of the aliquots are illustrated in Figure 2. This is the direct link between blood and tissue concentration needed for physiologic modeling. Even if the same well counter is used for both 2D and 3D measurements, errors in scanner calibration will influence the results.

Most of the tested devices deviated by <10%; nearly half were even better, deviating by <5%. Pronounced errors in accuracy, indicating possible malfunction of the instrument, were found in only a few devices (well counters 5, 12, 13, 16, 18, and 20). This group of instruments could be subdivided into 2 classes, those with errors in accuracy and those with, in addition, large variations between individual measurements (data not shown).

DISCUSSION

In quantitative PET, the image count rate per volume (representing tissue activity concentration) is related to the

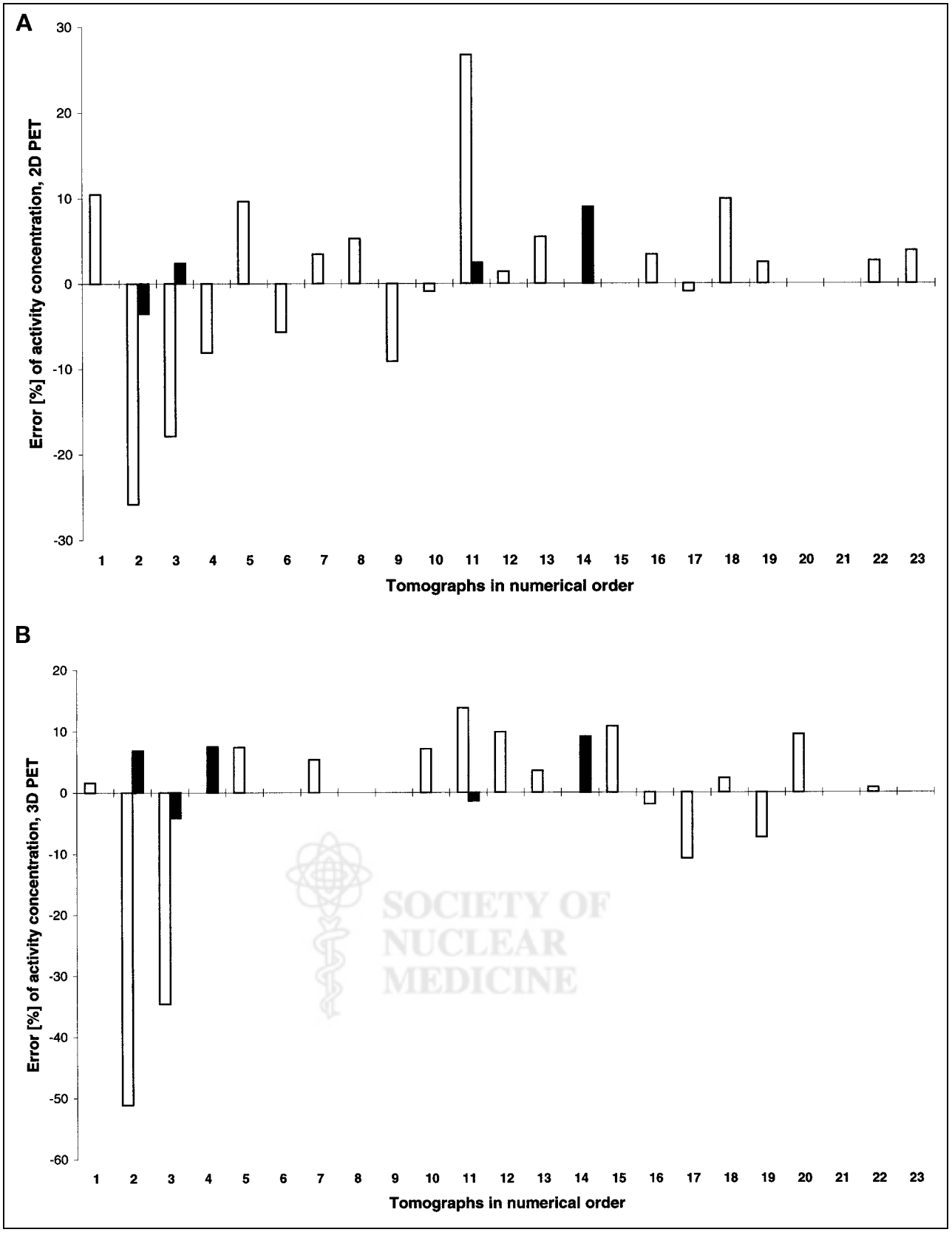

FIGURE 1. Percentage error of activity concentrations derived from PET images (white columns) for 2D (A) and 3D (B) acquisitions. Reference value is activity concentration as determined from dose calibrator measurement of activity injected and phantom volume. Black columns represent percentage errors after maintenance or corrective procedures. Scanners 6, 8, 9, and 23 could be operated in 2D mode only, and scanners 15, 20, and 21 could be operated in 3D mode only. Values of scanner 21 could not be evaluated. Initial values of scanner 14 for 2D and 3D and scanner 4 for 3D were not available.

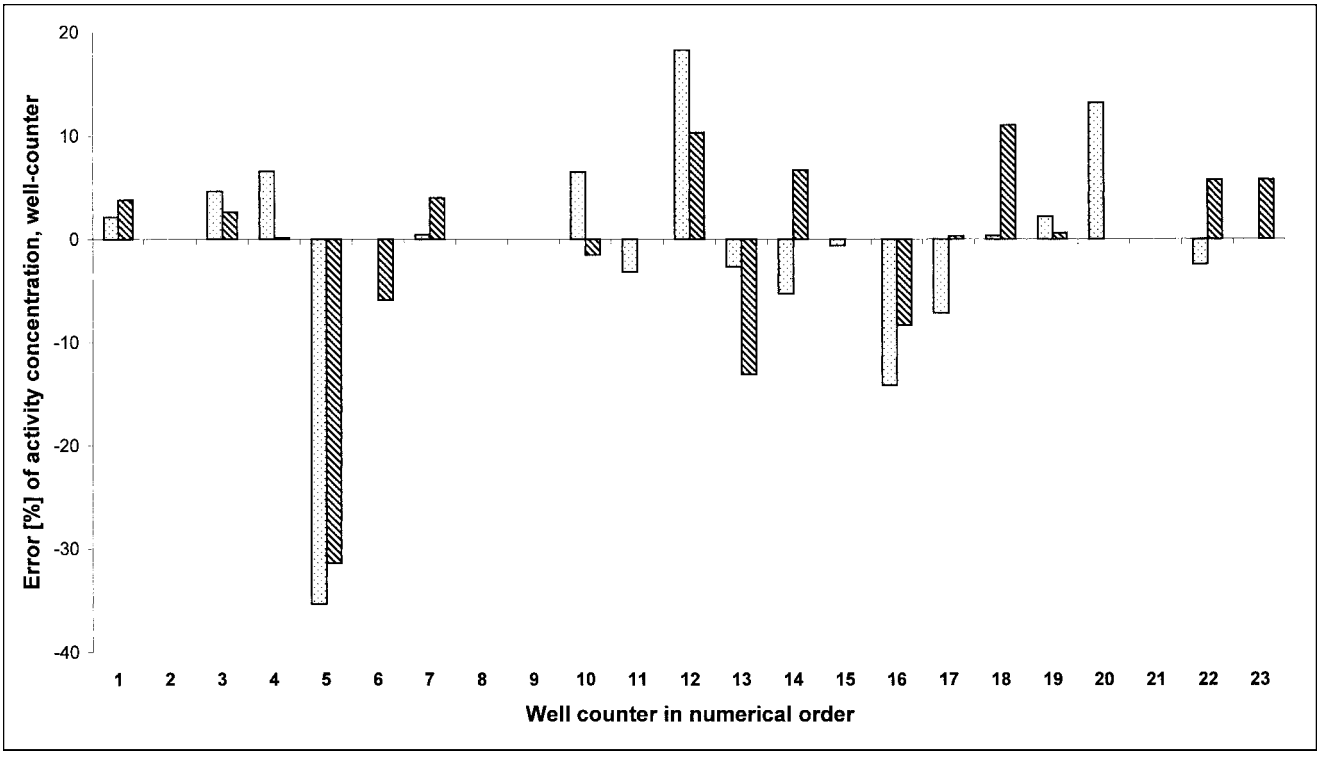

FIGURE 2. Mean percentage error of activity concentrations derived from PET images with respect to values measured in well counters for 2D (hatched columns) and 3D (dotted columns) modes. Respective data for well counters 2, 9, and 21 are missing, because well counters were not available at these centers. Error of well counter 8 is near zero and cannot be seen on this scale.

injected dose (e.g., determination of standardized uptake values) and to the blood activity concentration (physiologic modeling). Therefore, careful cross calibration between the PET scanner, dose calibrator, and well counter is essential.

The calibration accuracy of the scanners in this study was examined by comparing phantom data between institutions. When the accuracy was found to be degraded, we tried to identify the source of the error. We do not describe a new method of calibrating PET scanners, but we investigated the accuracy of the calibration process as performed in the field and in the context of a multicenter study. To facilitate analysis of errors and pooling of data from different institutions, we also checked the absolute accuracy of the dose calibrators using a set of certified standards. Nevertheless, the tests that could be performed in the framework of this study cover only the basic prerequisites for quantifying PET activity concentrations in vivo (*5*).

All investigated PET scanners were suitable for visual analysis of data, as judged by daily quality control and visual inspection of sinograms and reconstructed images. However, some scanners could not be used for quantitative studies without prior corrective maintenance. Quality control, although a prerequisite for operation of PET scanners, does not guarantee the accuracy of scanner calibration.

The calibration chain starts with the dose calibrator, whose accuracy is therefore of fundamental importance. In this respect, the test outcome was quite satisfying. Nearly all instruments had a $<10\%$ deviation, and most were even better ($<5\%$). The outliers were either defective because of age or contaminated—problems mandatory to be solved. Less satisfactory were the tests on scanner calibration and cross calibration to the dose calibrator. At the first attempt, only 50% of the devices showed $<5\%$ deviation, but after the problems had been identified and solved, the accuracy of most of the remaining devices could be improved to $\leq 10\%$. Only 1 scanner could not be fixed at all, because of unresolvable problems with hardware and software.

The sources of error ranged widely—from more sophisticated problems, such as misalignment of processor clocks and loss of some components of the normalization and correction software data, to simple user handling errors (these being the majority), such as failure to perform calibration at all, faulty specification of the phantom activity concentration, and errors in performing the calibration and cross calibration.

Only a few of the centers actually used measurement of blood samples as an input function for physiologic modeling and, therefore, needed a cross-calibrated well counter. Nevertheless, when this type of device was available it was checked. So the relatively large errors for some of the tested instruments may be explained by infrequent use and, therefore, no need for cross calibration.

CONCLUSION

For clinical multicenter studies relying on quantitative analysis of PET data, the calibration of scanners must

carefully be checked beforehand. By thorough application of straightforward, standard procedures, an accuracy of at least 5%–10% could be achieved for nearly all the dedicated PET scanners tested.

However, the measurements and readjustments performed during this study reflected the status of the devices at that time. Permanent monitoring by the participating institutions is needed to maintain this condition.

ACKNOWLEDGMENTS

The authors are grateful to Olaf Scheibe (AEA Technology QSA GmbH) for expert logistic support with the ^{68}Ge standards and to the staff of the participating institutions for help with measurements. This study was supported by the Deutsche Krebshilfe (grant 70-2329).

REFERENCES

1. Bailey DL. Quantitative procedures in 3D PET. In: Bendriem B, Townsend DW, eds. *The Theory and Practice of 3D PET*. Boston, MA: Kluwer Academic Publishers; 1998:55–109.
2. Eichling JO, Higgins CS, Ter-Pogossian MM. Determination of radionuclide concentrations with positron CT scanning (PETT): concise communication. *J Nucl Med.* 1977;18:845–847.
3. Bailey DL, Jones T. A method for calibrating three-dimensional positron emission tomography without scatter correction. *Eur J Nucl Med.* 1997;24:660–664.
4. Defrise M, Geissbuhler A, Townsend DW. A performance study of 3D reconstruction algorithms for positron emission tomography. *Phys Med Biol*. 1994;39: 305–320.
5. Knoop BO. PET: performance parameters and quantitative corrections. *Physica Medica.* 1992;8(suppl 1):59–67.