Multicenter Harmonization of ^{89}Zr PET/CT Performance

Nikolaos E. Makris[1], Ronald Boellaard[1], Eric P. Visser[2], Johan R. de Jong[3], Bruno Vanderlinden[4], Roel Wierts[5], Berlinda J. van der Veen[6], Henri J.N.M. Greuter[1], Danielle J. Vugts[1], Guus A.M.S. van Dongen[1], Adriaan A. Lammertsma[1], and Marc C. Huisman[1]

[1]Department of Radiology and Nuclear Medicine, VU University Medical Center, Amsterdam, The Netherlands; [2]Department of Nuclear Medicine, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands; [3]Department of Nuclear Medicine and Molecular Imaging, University Medical Center Groningen, Groningen, The Netherlands; [4]Department of Nuclear Medicine, Jules Bordet Institute, Brussels, Belgium; [5]Department of Nuclear Medicine, Maastricht University Medical Center, Maastricht, The Netherlands; and [6]Department of Nuclear Medicine, The Netherlands Cancer Institute, Amsterdam, The Netherlands

This study investigated the feasibility of quantitative accuracy and harmonized image quality in ^{89}Zr-PET/CT multicenter studies. **Methods:** Five PET/CT scanners from 3 vendors were included. ^{89}Zr activity was measured in a central dose calibrator before delivery. Local activity assays were based on volume as well as on the local dose calibrator. Accuracy and image noise were determined from a cross calibration experiment. Image quality was assessed from recovery coefficients derived from different volume-of-interest (VOI) methods ($VOI_{A50\%}$, based on a 3-dimensional isocontour at 50% of the maximum voxel value with local background correction; VOI_{Max}, based on the voxel with the highest uptake; and VOI_{3Dpeak}, based on a spheric VOI of 1.2-cm diameter positioned so as to maximize the enclosed average). PET images were analyzed before and after postreconstruction smoothing, applied to match image noise. **Results:** PET/CT accuracy and image noise ranged from −3% to 10% and from 13% to 22%, respectively. VOI_{3Dpeak} produced the most reproducible recovery coefficients. After calibration of the local dose calibrator to the central dose calibrator, differences between the local activity assays were within 6%. **Conclusion:** This study showed that quantitative accuracy and harmonized image quality can be reached in ^{89}Zr PET/CT multicenter studies.

Key Words: ^{89}Zr-PET; multi-center; image quality; cross calibration; harmonization

J Nucl Med 2014; 55:264–267
DOI: 10.2967/jnumed.113.130112

PET using labeled monoclonal antibodies, also known as immuno-PET, shows promise as a tool to predict the outcome of cancer treatment based on monoclonal antibodies (*1*). Their kinetics dictate the need for a positron label with a long half-life. An ideal radionuclide is ^{89}Zr (half-life, 78.41 h) because its physical half-life matches the biologic half-life of most antibodies. Additionally, it can easily and stably be coupled to monoclonal antibodies (*2*). To date, all ^{89}Zr-monoclonal antibody PET/CT studies that have been reported were performed within a single center (*3–5*). More recently, several multicenter studies have been initiated. Multicenter studies with ^{18}F-labeled tracers have shown the need for standardization of image acquisition, reconstruction, and analysis procedures, such as outlined in the European Association of Nuclear Medicine guidelines for tumor imaging (*6*) and implemented in the form of an accreditation (EANM Research Ltd. [EARL]). For ^{89}Zr, there are several additional factors that need to be considered: nonprompt emission of 909-keV γ rays after each positron emission may influence cross calibration between the local dose calibrator and the PET/CT camera, and low counting rates (due to both low positron abundance and low injected activity) may potentially lead to poorer image quality. The aim of this study was therefore to investigate the feasibility of quantitative accuracy and harmonized image quality in ^{89}Zr-PET/CT multicenter studies.

MATERIALS AND METHODS

Scanners

Three Gemini TF PET/CT scanners (Philips Healthcare) (*7*), a Biograph mCT PET/CT scanner (Siemens Medical Solutions) (*8*), and a Discovery-690 PET/CT scanner (GE Healthcare) (*9*) were used in this study. Four of 5 PET/CT scanners were EARL-accredited before the start of this study.

^{89}Zr Activity Concentration Measurements

After production of ^{89}Zr, a nonsticking solution was prepared for the phantom experiments. The solution consisted of 1 M oxalic acid neutralized with 2 M Na_2CO_3, diluted with 0.2 M 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid in phosphate buffer saline containing 1 mg·mL^{-1} of bovine serum albumin. All vials were measured in the central dose calibrator before delivery to the various PET centers, as well as in each local dose calibrator afterward. Assuming a homogeneous ^{89}Zr solution, injected activity in each phantom compartment was determined on the basis of measured net injected volume (derived from measurements of the weights of the syringes before and after injecting activity in the phantoms). In addition, local dose calibrator syringe measurements were performed to determine injected activity in each phantom compartment and compared with volume-based activity measurements.

Phantoms

A custom-made, homogeneous cylindric phantom (inner diameter, 29 cm; inner length, 20 cm) with a volume of 13.2 L was filled with an ^{89}Zr solution with an activity concentration of about 1 kBq·mL^{-1}. This phantom will be referred to as the cross calibration phantom. The 9.7-L background compartment of a National Electrical Manufacturers

Received Jul. 31, 2013; revision accepted Oct. 24, 2013.
For correspondence or reprints contact: Nikolaos E. Makris, VU University Medical Center, Department of Radiology and Nuclear Medicine, De Boelelaan 1117, 1081 HV Amsterdam, The Netherlands.
E-mail: n.makris@vumc.nl
Published online Dec. 19, 2013.

TABLE 1
PET/CT Reconstruction Settings

Scanner type	Reconstruction algorithm	Iterations	Subsets	Default smoothing (mm)	Sensitivity (cps/kBq)	Axial field of view (cm)	Axial pixel size (mm)	Additional smoothing (mm)
mCT	PSF-TOF	3	21	8.0	9.6	21.8	2.4	—
Gemini	BLOB-OS-TF	3	33	—	6.6	18.0	4.0	7
Discovery	OSEM-TOF	2	18	6.4	7.5	15.7	3.3	6

PSF-TOF = point spread function time of flight; BLOB-OS-TF = BLOB (rotationally symmetric volume elements) ordered-subsets time of flight; OSEM-TOF = ordered-subsets expectation maximization time of flight.

Association NU-2 image quality phantom (Data Spectrum) was filled with 0.7 kBq·mL^{-1}. This phantom contains 6 spheres with inner diameters of 10, 13, 17, 22, 28, and 37 mm. All spheres were filled with a sphere-to-background activity concentration of 10:1. This phantom will be referred to as the image quality phantom.

Acquisition and Reconstruction Protocols

A 10-min-per-bed-position 1-bed-position acquisition and a 5-min-per-bed-position 2-bed-position acquisition were obtained for the cross calibration and image quality phantoms, respectively. Data were normalized; corrected for decay, randoms, dead time, scatter, and attenuation; and reconstructed using settings (Table 1) associated with EARL accreditation.

The results of the present study were obtained on EARL-accredited scanners, as this accreditation program provides the most detailed specifications for harmonized quantitative performance (*10*). All issues addressed in the present study, however, should also be applicable to other accreditation programs, as they are not fundamentally dependent on the specific accreditation program being followed.

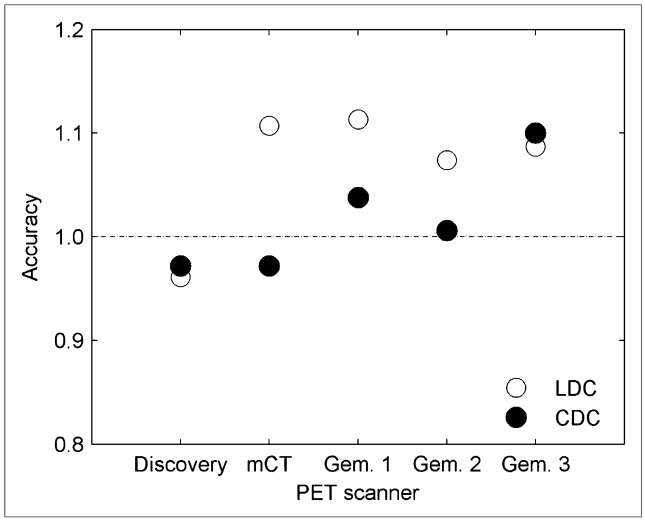

FIGURE 1. Accuracy of 5 different PET scanners relative to local dose calibrator (LDC) and central dose calibrator (CDC) measurements.

Data Analysis

Data were analyzed using software developed in-house. PET/CT calibration accuracy was defined as PET-measured activity divided by the activity measured in the central dose calibrator. Noise (%) was characterized by means of the coefficient of variation for a volume of interest (VOI) minimally 1 cm from the edge of the cross calibration phantom. Additional smoothing (Table 1) was applied to match image noise. The image quality phantom was analyzed by calculating the recovery coefficient (RC) as a function of sphere size, being defined as the ratio of PET activity concentration to central dose calibrator activity concentration. $VOI_{A50\%}$ was defined as VOI based on a 3-dimensional isocontour at 50% of the maximum voxel value with local background correction, VOI_{Max} was defined as VOI based on the voxel with the highest uptake, and VOI_{3Dpeak} was defined as VOI based on a spheric 1.2-cm-diameter VOI positioned so as to maximize the enclosed average (*11*).

RESULTS

Local dose calibrator activity measurements on the vials differed by up to 14% from central dose calibrator activity measurements. After correction for these differences, volume-based measurements coincided with syringe-based ones within 6%. In Figure 1, the accuracy of each PET scanner relative to local dose calibrator and central dose calibrator measurements is shown. Based on central dose calibrator measurements, both the Discovery and the mCT scanners showed an accuracy within 3%, and the 3 Gemini scanners showed accuracies of 4%, 1%, and 10%. Figure 2A shows that noise along the axial direction of the scans ranged from 13% to 25% in the center. Images from the Gemini and Discovery systems were smoothed (Table 1), and a comparable coefficient of variation of about 13% in the center resulted for all images (Fig. 2B).

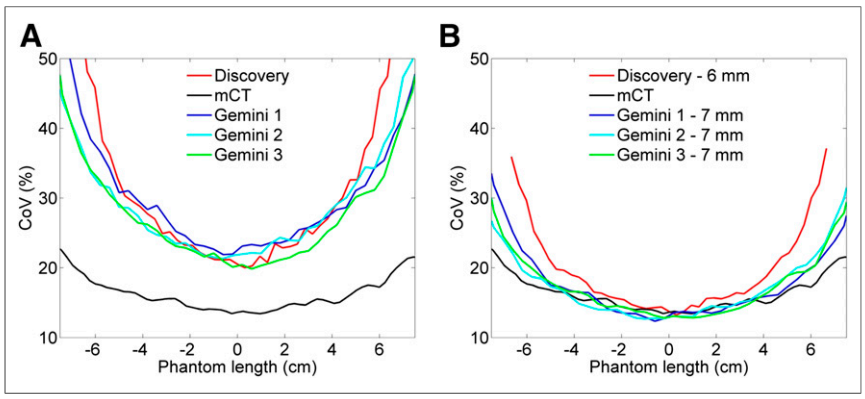

FIGURE 2. Coefficient of variation (CoV) per phantom unit length without (A) and with (B) additional smoothing data for 5 different scanners.

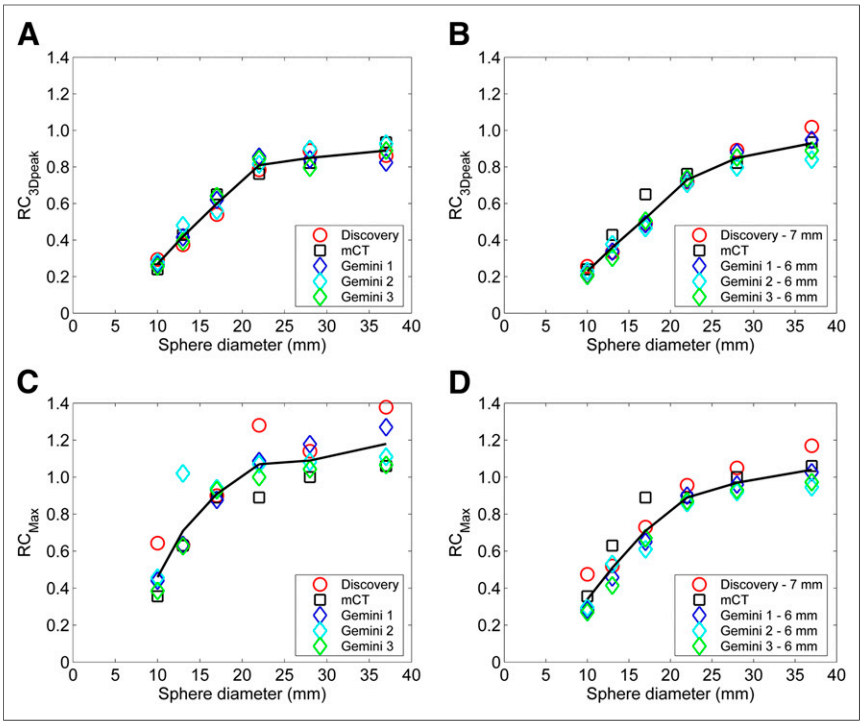

FIGURE 3. RC as function of sphere diameter for VOI_{3Dpeak} (A and B) and VOI_{Max} (C and D) without (A and C) and with (B and D) additional smoothing.

Figure 3 illustrates RCs for all PET scanners without and with the same additional smoothing, showing the lowest and largest RC variability for VOI_{3Dpeak} and VOI_{max}, respectively. Figure 4 shows average RC as a function of sphere diameter for the 3 VOI definition methods without and with additional smoothing.

DISCUSSION

This study assessed accuracy, noise, and RC characteristics for ^{89}Zr PET imaging with the goal of achieving accuracy and harmonized image quality in a multicenter setting.

Multicenter ^{89}Zr imaging essentially entails 3 additional steps for EARL-accredited scanners: calibration of the local dose calibrator for ^{89}Zr with respect to a common central dose calibrator; postreconstruction smoothing of images derived from a Gemini or Discovery system; and use of VOI_{3Dpeak} for analysis of activity concentrations in lesions.

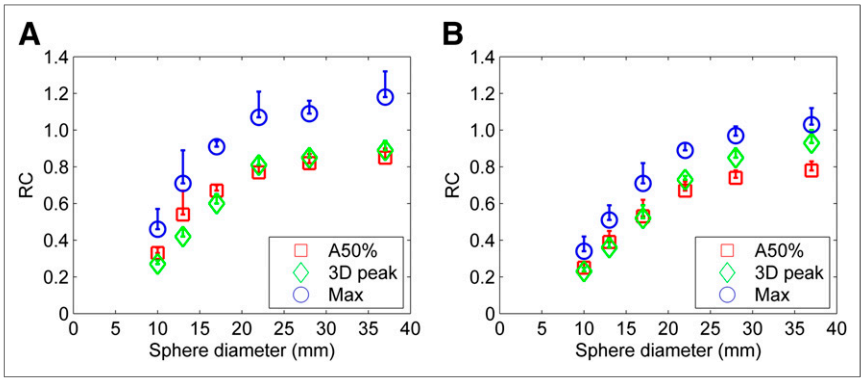

FIGURE 4. Average RC as function of sphere diameter without (A) and with (B) additional smoothing data for 5 different scanners.

A limitation of the present study was the small sample size. However, information on image quality harmonization between scanners and institutions is already available (*6*) and has led to the EARL accreditation program, which aims to reduce this variation specifically for whole-body ^{18}F-FDG scanning in an oncology setting. Therefore, only additional issues specific to ^{89}Zr needed to be addressed. On the basis of existing EARL accreditation criteria, it seemed sufficient to include one current-generation scanner from each of the main vendors and to perform a small number of repeated measurements on a scanner of a single vendor, as the impact of low counting statistics and nonprompt γ radiation may be assumed to be similar for similar types of PET/CT systems.

Cross calibration accuracy between the central dose calibrator and the PET/CT systems ranged from −3% to +10%, whereas use of individual uncalibrated local dose calibrators increased the variability (Fig. 1). The largest inaccuracy (+10%) was shown for the non–EARL-accredited PET/CT scanner (Gemini 3). After calibration of the local dose calibrator to a central dose calibrator, volume-based syringe activity assessments agreed within 6% with local dose calibrator measurements. Therefore, clinical work can be based on local dose calibrator measurements, although for multicenter clinical trials, calibration of all local dose calibrators against a central dose calibrator is recommended. Noise levels ranged from 13% to 22% (central plane). The mCT showed the lowest coefficient of variation and a less curved profile. Because parameters such as injected activity, number of bed positions, and time per bed position were the same in all scanners, this observation may be attributed to the extended field of view of the mCT scanner and its associated higher sensitivity (Table 1). Additional smoothing of the images from the other PET/CT systems was applied to obtain comparable noise levels in the central area of the cross calibration phantom (Fig. 2). Because additional smoothing can downgrade image resolution, the effect of smoothing on RC was also investigated.

The use of VOI_{3Dpeak} resulted in a somewhat larger RC range than that obtained with $VOI_{A50\%}$. This can be explained, at least in part, by using a 1.2-cm fixed-size diameter for VOI_{3Dpeak}. For the largest spheres, RC with $VOI_{A50\%}$ is based on larger volumes than is RC with VOI_{3Dpeak}, resulting in lower RC. The variability of RC among PET scanners was more prominent with VOI_{Max} because of the use of a single voxel for determining RC, associated with higher noise levels. A significantly lower RC variability was observed for $VOI_{A50\%}$ and VOI_{3Dpeak}. A similar trend was observed for the additionally smoothed data. These observations are consistent with a previous study for ^{18}F-FDG

(*10*). Another study (*12*) has shown that a large fixed-size square VOI may result in a poor estimate of standardized uptake value (SUV) response but only when metabolic tumor sizes decrease below 1 mL. Yet, a recent clinical study reported low variability for peak SUV (*13*). Moreover, variability of RC based on VOI_{3Dpeak} was the smallest among VOI methods, consistent with a report (*14*) suggesting that VOI_{3Dpeak} may be more robust to changes in pixel size and image characteristics. SUV_{3Dpeak} may therefore be preferable for use in ^{89}Zr multicenter studies.

CONCLUSION

This study investigated the potential for performing quantitative multicenter ^{89}Zr PET/CT studies assuming harmonized data acquisition and reconstruction settings. After recalibration of the local dose calibrator to the central dose calibrator for ^{89}Zr, local activity measurements were accurate within 6%. After matching noise levels (~13% at the center of the cross calibration phantom), the use of VOI_{3Dpeak} resulted in ±7% variability in RC for each sphere across various PET/CT systems and imaging sites. The use of an ^{89}Zr calibration procedure in combination with SUV_{3Dpeak} in image analysis is recommended to be able to perform quantitatively accurate multicenter ^{89}Zr PET/CT studies with a harmonized image quality.

DISCLOSURE

The costs of publication of this article were defrayed in part by the payment of page charges. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734. The study was financially supported in part by Philips Healthcare. No other potential conflict of interest relevant to this article was reported.

REFERENCES

1. Wu AM. Antibodies and antimatter: the resurgence of immuno-PET. *J Nucl Med.* 2009;50:2–5.
2. Verel I, Visser GW, Boellaard R, Stigter-van WM, Snow GB, van Dongen GA. ^{89}Zr immuno-PET: comprehensive procedures for the production of ^{89}Zr-labeled monoclonal antibodies. *J Nucl Med.* 2003;44:1271–1281.
3. Börjesson PK, Jauw YW, de Bree R, et al. Radiation dosimetry of ^{89}Zr-labeled chimeric monoclonal antibody U36 as used for immuno-PET in head and neck cancer patients. *J Nucl Med.* 2009;50:1828–1836.
4. Dijkers EC, Oude Munnink TH, Kosterink JG, et al. Biodistribution of ^{89}Zr-trastuzumab and PET imaging of HER2-positive lesions in patients with metastatic breast cancer. *Clin Pharmacol Ther.* 2010;87:586–592.
5. Rizvi SN, Visser OJ, Vosjan MJ, et al. Biodistribution, radiation dosimetry and scouting of ^{90}Y-ibritumomab tiuxetan therapy in patients with relapsed B-cell non-Hodgkin's lymphoma using ^{89}Zr-ibritumomab tiuxetan and PET. *Eur J Nucl Med Mol Imaging.* 2012;39:512–520.
6. Boellaard R, O'Doherty MJ, Weber WA, et al. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging—version 1.0. *Eur J Nucl Med Mol Imaging.* 2010;37:181–200.
7. Surti S, Kuhn A, Werner ME, Perkins AE, Kolthammer J, Karp JS. Performance of Philips Gemini TF PET/CT scanner with special consideration for its time-of-flight imaging capabilities. *J Nucl Med.* 2007;48:471–480.
8. Jakoby BW, Bercier Y, Conti M, Casey ME, Bendriem B, Townsend DW. Physical and clinical performance of the mCT time-of-flight PET/CT scanner. *Phys Med Biol.* 2011;56:2375–2389.
9. Bettinardi V, Presotto L, Rapisarda E, Picchio M, Gianolli L, Gilardi MC. Physical performance of the new hybrid PETCT Discovery-690. *Med Phys.* 2011;38:5394–5411.
10. Makris NE, Huisman MC, Kinahan PE, Lammertsma AA, Boellaard R. Evaluation of strategies towards harmonization of FDG PET/CT studies in multicentre trials: comparison of scanner validation phantoms and data analysis procedures. *Eur J Nucl Med Mol Imaging.* 2013;40:1507–1515.
11. Cheebsumon P, Yaqub M, van Velden FH, Hoekstra OS, Lammertsma AA, Boellaard R. Impact of [^{18}F]FDG PET imaging parameters on automatic tumour delineation: need for improved tumour delineation methodology. *Eur J Nucl Med Mol Imaging.* 2011;38:2136–2144.
12. Boellaard R, Krak NC, Hoekstra OS, Lammertsma AA. Effects of noise, image resolution, and ROI definition on the accuracy of standard uptake values: a simulation study. *J Nucl Med.* 2004;45:1519–1527.
13. Benz MR, Allen-Auerbach MS, Eilber FC, et al. Combined assessment of metabolic and volumetric changes for assessment of tumor response in patients with soft-tissue sarcomas. *J Nucl Med.* 2008;49:1579–1584.
14. Lodge MA, Chaudhry MA, Wahl RL. Noise considerations for PET quantification using maximum and peak standardized uptake value. *J Nucl Med.* 2012;53:1041–1047.