Multicentre harmonization of $^{89}$Zr PET/CT performance


Abstract

This study investigated the feasibility of quantitative accuracy and harmonized image quality in $^{89}\text{Zr}$-PET/CT multicentre studies.

Methods: Five PET/CT scanners from 3 vendors were included. $^{89}\text{Zr}$ 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 ($\text{VOI}_{50\%}$, based on a 3-dimensional isocontour at 50% of the maximum voxel value with local background correction; $\text{VOI}_{\text{max}}$, based on the voxel with the highest uptake; and $\text{VOI}_{3D\text{peak}}$, based on a spheric VOI of 1.2 cm diameter positioned so as to maximize the enclosed average). PET images were analysed before and after post-reconstruction smoothing, applied to match image noise.

Results: PET/CT accuracy and image noise ranged from -3% to 10% and from 13% to 22%, respectively. $\text{VOI}_{3D\text{peak}}$ 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 $^{89}\text{Zr}$ PET/CT multicentre studies.
3.1 Introduction

PET using labeled monoclonal antibodies, also known as immunoPET, 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 centre [3-5]. More recently, several multicentre studies have been initiated. Multicentre 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 tumour 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 $\gamma$ rays after each positron emission may influence cross calibration between the local dose calibrator and the PET/CT camera, and low counting 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 multicentre studies.

3.2 Materials and methods

3.2.1 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.

3.2.2 $^{89}$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$_2$CO$_3$, 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 centres, as well as in each local dose calibrator afterward.
Chapter 3. Multicentre harmonization of $^{89}$Zr PET/CT performance

Table 3.1. PET/CT reconstruction settings; PSF-TF = point spread function time of flight; BLOB-OS-TF = BLOB (rotationally symmetric volume elements) ordered-subsets time of flight; OSEM-TF = ordered-subsets expectation maximization time of flight.

<table>
<thead>
<tr>
<th>Scanner type</th>
<th>Reconstruction algorithm</th>
<th>Iterations</th>
<th>Subsets</th>
<th>Default Sensitivity (cps/kBq)</th>
<th>Axial field (mm)</th>
<th>Axial pixel size (mm)</th>
<th>Additional smoothing (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mCT</td>
<td>PSF-TF</td>
<td>3</td>
<td>21</td>
<td>8.0</td>
<td>21.8</td>
<td>2.4</td>
<td>-</td>
</tr>
<tr>
<td>Gemini</td>
<td>BLOB-OS-TF</td>
<td>3</td>
<td>33</td>
<td>-</td>
<td>18.0</td>
<td>4.0</td>
<td>7</td>
</tr>
<tr>
<td>Discovery</td>
<td>OSEM-TF</td>
<td>2</td>
<td>18</td>
<td>6.4</td>
<td>15.7</td>
<td>3.3</td>
<td>6</td>
</tr>
</tbody>
</table>

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.

3.2.3 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 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.

3.2.4 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 3.1) 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.

### 3.2.5 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 3.1) 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$_{450\%}$ 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 (II).

### 3.3 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 3.1, the accuracy of each PET scanner
Chapter 3. Multicentre harmonization of $^{89}$Zr PET/CT performance

Figure 3.2. Coefficient of variation (CoV) per phantom unit length without (A) and with (B) additional smoothing data for 5 different scanners.

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 3.2A shows that noise along the axial direction of the scans ranged from 13% to 25% in the centre. Images from the Gemini and Discovery systems were smoothed (Table 3.1), and a comparable coefficient of variation of about 13% in the centre resulted for all images (Fig. 3.2B).

Figure 3.3 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 3.4 shows average RC as a function of sphere diameter for the three VOI definition methods without and with additional smoothing.

3.4 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 multicentre setting.

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

A limitation of the present study was the small sample size. However, information on image quality harmonization between scanners and institu-
3.4. Discussion

Figure 3.3. RC as function of sphere diameter for VOI$_{\text{Dpeak}}$ (A and B) and VOI$_{\text{Max}}$ (C and D) without (A and C) and with (B and D) additional smoothing.

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. 3.1). The largest inaccuracy (+10%) was shown for the nonEARL-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 multicentre 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 3.1). 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. 3.2). Because additional smoothing can downgrade image resolution, the effect of smoothing on RC was also investigated.

The use of VOI$\text{3D}_{\text{peak}}$ resulted in a somewhat larger RC range than that obtained with VOI$\text{A}_{50\%}$. This can be explained, at least in part, by using a 1.2-cm fixed-size diameter for VOI$\text{3D}_{\text{peak}}$. For the largest spheres, RC with VOI$\text{A}_{50\%}$ is based on larger volumes than is RC with VOI$\text{3D}_{\text{peak}}$, resulting in lower RC. The variability of RC among PET scanners was more prominent with VOI$\text{Max}$ 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$\text{A}_{50\%}$ and VOI$\text{3D}_{\text{peak}}$. 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 tumour sizes decrease below 1 mL. Yet, a recent clinical study reported low variability for peak SUV [13].

Figure 3.4. Average RC as function of sphere diameter without (A) and with (B) additional smoothing data for 5 different scanners.
Moreover, variability of RC based on VOI_{3D\text{peak}} was the smallest among VOI methods, consistent with a report [14] suggesting that VOI_{3D\text{peak}} may be more robust to changes in pixel size and image characteristics. SUV_{3D\text{peak}} may therefore be preferable for use in $^{89}$Zr multicentre studies.

3.5 Conclusions

This study investigated the potential for performing quantitative multicentre $^{89}$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 ($\sim 13\%$ at the centre of the cross calibration phantom), the use of VOI_{3D\text{peak}} resulted in $\pm 7\%$ 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_{3D\text{peak}} in image analysis is recommended to be able to perform quantitatively accurate multicentre $^{89}$Zr PET/CT studies with a harmonized image quality.
References
