Multicenter integrated positron emission tomography/computer tomography imaging quality assurance accreditation interobserver reliability study with the American College of Radiology phantom

Pan-Fu Kao, Kuo-Wei Lee, Chi-Wen Huang, Shun-Fa Yang, Thomas Chang-Yao Tsao

Abstract

Background: Integrated positron emission tomography/computer tomography (PET/CT) image quality assurance ensures accurate, reproducible, and quantitative assessment of comparable scanner performance. We performed a preliminary multicenter PET/CT imaging quality assurance test with a fillable tomographic phantom in six medical centers in Taiwan.

Methods: The phantom was filled with fixed proportions of fluorine-18 radionuclide solution in the background and with different spheres to simulate cold and hot lesions, and body background radioactivity. Imaging acquisitions were performed by using recommended parameters in different sites according to different brand names of the instrument. All imaging was subjectively scored by eight experienced nuclear medicine physicians as the spatial resolution of hot vials (score 0-4), six cold spheres (score 0-6), and six cold rod areas (score 0-6), and overall satisfaction (score 0-5). Interobserver correlation and receiver operating characteristic (ROC) curve were analyzed.

Results: The detection ability of hot vials, cold spheres, and cold rods was 4.0 ± 0.1, 5.2 ± 0.8, and 3.8 ± 0.9, respectively. Overall satisfaction was 4.0 ± 0.8. The ROC analysis revealed that the area under the curve for hot vials, cold spheres, and cold rods was 0.984, 0.887, and 0.928 respectively. The interobserver correlation for detectability of cold spheres and cold rods was 0.88 and 0.96, respectively.

Conclusion: The results of the study indicated that (1) PET/CT imaging quality assurance for comparable scanner performance could be established on the basis of a standard phantom and (2) good interobserver correlation can be observed for those with accurate and interpretable results.

Keywords: Fluorine-18; Integrated positron emission tomography/computer tomography (PET/CT); Quality assurance; Tomographic phantom

1. INTRODUCTION

Fluorine-18-labeled radiopharmaceuticals, especially 2-[fluorine-18]-fluoro-2-deoxy-d-glucose ([18F-FDG]), positron emission tomography (PET) images, provide molecular-level information noninvasively. PET images can be used for diagnosis, staging, and therapy monitoring of different cancers, myocardial viability, and brain disease. The development of integrated PET/computed tomography (CT) system provides not only in vivo functional information but also accurate anatomical localization and attenuation correction.

In 2006, MacFarlane reported the status of accreditation of nuclear medicine and PET imaging departments in the United States. The accreditation program of the American College of Radiology (ACR) for PET accreditation required a PET performance test with a phantom image. The most important goal of ACR for PET accreditation is to improve patient care. In 2010, the European Association of Nuclear Medicine launched an [18F-FDG PET/CT accreditation program. The aim of the program is to help imaging sites meet the standard requirements, ensuring accurate and reproducible quantitative assessment of comparable scanner performance in all participants of multicenter trials. With the promotion of fluorine-18-labeled radiopharmaceuticals PET/CT imaging in clinical research and the large population of cancer patients in Taiwan, the nuclear medicine society in Taiwan has a strong intention on enhancing the quality of fluorine-18 PET/CT clinical practice and research to provide high-quality health care.
This preliminary study was to perform a multicenter accreditation PET performance test with an ACR Jaszczak Deluxe Flangeless phantom filled with fluorine-18-labeled radiopharmaceuticals to assure PET/CT imaging quality in Taiwan. The purpose of this study was to validate accredited members’ interobserver correlation and image interpretation consistency of spatial resolution and lesion detection.

2. METHODS

The study recruited six PET/CT scanners from six different medical centers in Taiwan to acquire an ACR-approved PET version of the tomographic phantom for PET/CT $^{18}$F radioisotope image quality assurance. The PET/CT scanners included four GE Healthcare Discovery series scanners, one Philips GXL scanner, and one Siemens mCT scanner.

2.1. Phantom preparation

A commercially available cylinder, the Jaszczak Deluxe Flangeless ECT phantom ACR-approved PET version of the tomographic phantom (Biodex Inc., Shirley, NY, USA), was used for the study to measure spatial resolution and the detectability of “hot” and “cold” lesions.

The ACR PET phantom is a cylinder with an internal diameter of 20.4 cm, containing three portions to insert testing apparatus for different purposes (Fig. 1A). The lower portion consisted of six sets of acrylic rods arranged in a six pie-shaped pattern (4.8, 6.4, 7.9, 9.5, 11.1, and 12.7 mm in diameter, respectively). The middle portion consisted of six cold spheres (9.5, 12.7, 15.9, 19.1, 25.4, and 31.8 mm in diameter, respectively). The upper faceplate with (1) four fillable thin-walled cylinders: 8, 12, 16, and 23 mm in diameter for radioactivity solution “hot” lesion and (2) three additional 25 mm cylinders: one for air, one for nonradioactive “cold” water, and one Teflon cylinder.

The phantom preparation followed the ACR phantom manual. To ensure standard uptake value (SUV) measurements for the PET scanning protocol, a 70 kg patient and the appropriate dose (e.g. 370 MBq ± 10%) was prepared, targeting a dose of 13.0 MBq in 1000 mL bag for hot vials activity dilution and a dose of 30.7MBq for background activity. Phantom activity was prepared right before each study on site. Scanning began 1 hour after the first vial activities were measured.

Imaging Acquisition and Reconstruction. The ACR phantom was placed on the image table properly with laser beam guidance for vertical and horizontal alignments (Fig. 1B). Imaging acquisitions were performed by using recommended parameters for routine clinical whole-body imaging protocol according to different brand names of the PET/CT scanners (Table 1). The tomography reconstruction parameters including iterations, subsets, and high- and low-pass filters were used on the basis of each manufactures’ recommended parameters to simulate daily clinical image conditions. The standards were based on results obtained from a variety of PET systems operating satisfactorily. The mean and SD of SUV in both cold and hot lesions were measured by putting circle region of interest (ROI) based on the outline of each spheres on CT images, and then copying and adjusting the ROI on the PET images.

Images Interpretation. An imaging review panel of eight readers was organized as the accreditation committee. At the time of the study, all readers had been specialty board-certified in nuclear medicine for >15 years. Each reader was considered to have an appropriate level of experience in the review of $^{18}$F-FDG PET/CT cases and the reporting of results without supervision. All PET/CT images acquired from six scanners were saved as digital imaging and communications in medicine (DICOM) files and were submitted to imaging review panel for scoring. Each of the readers was asked independently to evaluate spatial resolution of $^{18}$F PET/CT phantom images, including hot vials, cold spheres, and cold rods.

Spatial resolution was judged by identifying the smallest cold rods, cold spheres, and hot vials in the ACR-approved phantom. A numerical score system was applied for identifying the smallest hot vials (score 1-4), cold spheres (score 1-6), and cold rods (score 1-6) to test the interobserver correlation, receiver operating characteristic (ROC) curve analysis, and area under curve (AUC) analysis. To determine the overall satisfaction rates, readers were asked to give a global rating for each image about overall impression by a 5-point score system (score 1-5): 5 as excellent, 3 as acceptable, and 1 as not diagnosable.

2.2. Statistical analysis

All statistical analyses were performed by using commercially available software (SPSS version 18.0; SPSS Inc., Chicago, IL, USA). Interobserver agreement was calculated by using the
intraclass correlation coefficient (ICC) with 95% CI, which accounted for the correlation among multiple readers for cold rods, cold spheres, and hot vials interpretation.

3. RESULTS

The PET/CT imaging of the ACR PET phantom is demonstrated in Fig. 2. The mean and SD of the SUVs of each cold and hot sphere are listed in Table 2.

### Table 1

<table>
<thead>
<tr>
<th>Hospitals</th>
<th>H1</th>
<th>H2</th>
<th>H3</th>
<th>H4</th>
<th>H5</th>
<th>H6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acquisition protocol</td>
<td>Whole-body 3D</td>
<td>Whole-body 3D</td>
<td>Whole-body 3D</td>
<td>Whole-body 2D</td>
<td>Whole-body 3D</td>
<td>Whole-body 2D</td>
</tr>
<tr>
<td>Scan mode</td>
<td>2 min/bed</td>
<td>4 min/bed</td>
<td>2 min/bed</td>
<td>4 min/bed</td>
<td>4 min/bed</td>
<td>4 min/bed</td>
</tr>
<tr>
<td>Energy window</td>
<td>425-650 keV</td>
<td>435-650 keV</td>
<td>511 keV ± 7.5%</td>
<td>511 keV ± 10%</td>
<td>511 keV ± 10%</td>
<td>511 keV ± 10%</td>
</tr>
<tr>
<td>Matrix size</td>
<td>192 × 192</td>
<td>200 × 200</td>
<td>192 × 192</td>
<td>128 × 128</td>
<td>128 × 128</td>
<td>128 × 128</td>
</tr>
<tr>
<td>Reconstruction method</td>
<td>OSEM</td>
<td>OSEM</td>
<td>OSEM</td>
<td>OSEM</td>
<td>OSEM</td>
<td>FBP</td>
</tr>
<tr>
<td>Iteration</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>...</td>
</tr>
<tr>
<td>Subset</td>
<td>24</td>
<td>21</td>
<td>32</td>
<td>21</td>
<td>21</td>
<td>...</td>
</tr>
<tr>
<td>Attenuation correction</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Scatter correction</td>
<td>-</td>
<td>Y</td>
<td>-</td>
<td>Y</td>
<td>-</td>
<td>Y</td>
</tr>
<tr>
<td>Partial volume correction</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>Y</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>PSF correction</td>
<td>...</td>
<td>Y</td>
<td>...</td>
<td>Y</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

**Fig. 2** The computer tomography (CT; left), positron emission tomography (PET; middle), and PET/CT fusion (right) imaging results of the American College of Radiology (ACR) tomographic phantom from the H1 hospital. **A**, The hot vials were all well demonstrated. The three additional 25 mm cylinders were all demonstrated well, with best clear margin on Teflon cylinder and then water and air. **B**, The cold spheres were all resolved well except the smallest one 9.5 mm cold sphere with a low contrast. **C**, For the cold rods of pie-shaped patterns, the smallest 6.4 mm one was not resolved well, the next 7.9 mm one was resolved with a low contrast, and the 9.5 mm one was resolved well.

3.1. Interpretation of images

The results of PET/CT subjective measurements from PET/CT images are listed in Table 3. The hot vials are most clearly demonstrated with a high contrast to the smallest 8 mm diameter sphere, except for one with a low contrast of the H6 image. The cold spheres could be resolved to the 12.7 mm sphere in most images, except the H4 image, which could only demonstrate 15.9 mm. For cold rods, all the 9.5 mm rods were resolved with a low contrast. The overall satisfaction was 4.0 ± 0.8,
indicating all readers felt confident that all six hospitals will pass the $^{18}$F-labeled radiopharmaceuticals PET/CT imaging quality assurance.

According to ROC analysis, AUCs of hot vials, cold spheres, and cold rods were 0.984, 0.887, and 0.928, respectively (Fig. 3). The optimal cutoff point for hot vials, cold spheres, and cold rods was 3.5, 14.3, and 8.7 mm, respectively (Table 4). The ICCs of interobserver agreement rates for cold spheres and cold rods were 0.964 (95% CI: 0.89-0.99) and 0.880 (95% CI: 0.66-0.98), respectively.

4. DISCUSSION
In the last decade, the development of $^{18}$F-labeled radiopharmaceuticals PET/CT image in clinical practice and medical research has led to the quest of higher quality PET/CT image to benefit
public health. It is important to establish $^{18}$F-labeled radiopharmaceuticals PET/CT accreditation to achieve comparable scanner performance across multiple sites for patient and research image results to be compared, exchanged, and combined. In this study, the preliminary $^{18}$F-labeled radiopharmaceuticals PET/CT imaging multicenter quality assurance test was achieved in Taiwan with a standard phantom.

In this study, spatial resolution was evaluated by visual inspection on images. The detection of hot and cold objects was determined by readers’ evaluation, which was compatible with clinical imaging reading process. Good interobserver correlation among the imaging review panel of eight senior nuclear medicine physicians confirmed the phantom study to be a reliable accreditation tool for PET/CT imaging quality assurance. From the results of this preliminary multicenter study, the review panel recommended that the imaging quality assurance test must be performed at least semiannually to fulfill the worldwide requirements.

For hot vial detection ability, all the PET/CT scanners except H6 were able to detect the smallest 8.0 mm hot vial with a high contrast. The hot lesion detectability is especially important for $^{18}$F-FDG whole-body imaging for oncological indications. For ACR phantom test recommendation, the satisfactory image quality of hot vial should be 12 mm vial resolved with a low contrast and 16 mm vial resolved with a high contrast. All the PET/CT scanners from six hospitals fulfilled the detection ability of hot lesions.

Although detection ability of cold lesions is less important in current clinical $^{18}$F-FDG whole-body PET/CT imaging for oncological indications, the study also confirmed good detection ability and interobserver correlation for cold sphere and cold rod lesions. The cold lesion identification was important for other radiotracers of brain or cardiac studies, such as cerebral vascular disease and myocardial perfusion defect lesions.

The image resolution is affected by scattering photons and transmission correction. In this study, we performed phantom acquisition based on routine clinical whole-body imaging protocol according to different brand names of the PET/CT scanners. Therefore, the difference in image resolution may result from differences in these acquisition parameters, as well as hardware designed. By comparing this phantom study results, each PET/CT center was recommended to review and adjust its scanners’ acquisition parameters for better image quality. One of the major limitations of the study roots from different manufacturer types of PET/CT scanners’ hardware design, which resulted in different recommended imaging acquisition and reconstruction parameters. However, the main purpose of this preliminary accreditation of PET/CT imaging was to assure the image quality on the basis of daily clinical practice.

Fusion PET/CT image improves specificity and sensitivity for tumor assessment by integrating potential PET molecular images and CT anatomical images. The CT image was also used for PET imaging attenuation correction. Because PET and CT were executed separately, which might result in image misregistration. In previous studies, the maximum spatial error in PET-CT misregistration experiments on average was <4 mm, which is better than the spatial resolution of the clinical PET scanner used and is considered sufficient for most PET/CT applications. Although we did not specifically include the assessment of misregistration between PET and CT in this study, the PET/CT imaging interpretation for spatial resolution of hot vials and cold spheres was already included in the imaging degradation factors caused by misregistration.

In 2010, the National Cancer Institute of United States developed the Centers for Quantitative Imaging Excellence (CQIE), requiring PET/CT imaging qualification for all quantitative imaging-based oncological clinical trials. Over a period of 5 years, 65 unique PET/CT scanners across 56 sites were submitted for CQIE initial qualification; 64 scanners passed the qualification. Our preliminary study set a standard for fluorine-18 PET/CT image quality assurance test in Taiwan for all PET/CT scanners in clinical practice using visual lesion detection; six PET/CT scanners all passed the accreditation by eight readers.

There are some limitations in this study. First, the dose calibrators we used to prepare radiopharmaceuticals activity in this multicenter phantom study were not crossed calibrated before the study. However, the accuracy of dose calibrators in each center was validated annually by a standard source from the Institute of Nuclear Energy Research in Taiwan. Therefore, we assumed that the variation in dose measurement will be minor. Second, image acquisition and reconstruction parameter applied in this study were based on recommendations from PET/CT manufacturers. Although there are no identical standards among sites, we are convinced that the results are useful in representing the routine whole-body PET/CT image quality of clinical practice in Taiwan. Third, the study results only applied to $^{18}$F-isotope radiopharmaceuticals because different positron ranges from $^{11}$C, $^{11}$N, $^{68}$Ga, and other positron emitters may affect PET imaging quality. Clinical trials involving using different positron emitter-labeled radiopharmaceuticals in PET/CT imaging evaluation should perform separate phantom studies in addition to $^{18}$F-radioisotopes.

In conclusion, this phantom study provided every aspects of imaging quality assurance for PET/CT scanners in six medical centers of Taiwan. The study results indicated that PET/CT imaging quality accreditation based on a standard phantom could be established to compare scanner performance. Accurate interpretable results can be seen with good interobserver correlation. The quality assurance of PET/CT images should be performed at least at elective centers for routine clinical diagnosis and for all the centers involving clinical trials.

ACKNOWLEDGMENTS

This work was supported in part by the Institute of Nuclear Energy Research, Atomic Energy Council, Executive Yuan of Republic of China, Taiwan [NL-1040368] and the Ministry of Science and Technology of Republic of China, Taiwan [MOST 104-2623-E-040-001-NU]. The authors thank Yu-Shun Wang, MSc, for statistical analyses and Chiao-En Joanne Kao, MSc, for editing the article.

REFERENCES


Table 4

<table>
<thead>
<tr>
<th></th>
<th>AUC</th>
<th>95% CI</th>
<th>Cut point, mm</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hot vials</td>
<td>0.984</td>
<td>0.000-1.000</td>
<td>3.5</td>
<td>0.968</td>
<td>1.000</td>
</tr>
<tr>
<td>Cold spheres</td>
<td>0.987</td>
<td>0.844-0.930</td>
<td>1.43</td>
<td>0.762</td>
<td>0.925</td>
</tr>
<tr>
<td>Cold rods</td>
<td>0.928</td>
<td>0.899-0.958</td>
<td>8.7</td>
<td>0.771</td>
<td>0.920</td>
</tr>
</tbody>
</table>

AUC = area under curve; CI = confidence interval.


