

## Japanese guideline for the oncology FDG-PET/CT data acquisition protocol: synopsis of Version 2.0

Hiroyoshi Fukukita · Kazufumi Suzuki · Keiichi Matsumoto ·  
Takashi Terauchi · Hiromitsu Daisaki · Yasuhiko Ikari ·  
Naoki Shimada · Michio Senda

Received: 1 April 2014 / Accepted: 1 April 2014 / Published online: 24 May 2014  
© The Japanese Society of Nuclear Medicine 2014

**Abstract** This synopsis outlines the Japanese guideline Version 2.0 for the data acquisition protocol of oncology FDG-PET/CT scans that was created by a joint task force of the Japanese Society of Nuclear Medicine Technology, the Japanese Society of Nuclear Medicine and the Japanese Council of PET Imaging, and was published in *Kakuigaku-Gijutsu* 2013; 33:377–420 in Japanese. The guideline aims at standardizing the PET image quality among PET centers and different PET camera models by providing criteria for the IEC body phantom image quality as well as for the patient PET image quality based on the noise equivalent count (NEC), NEC density and liver signal-to-noise ratio, so that the appropriate scanning parameters can be determined for each PET camera. This Version 2.0 covers issues that were not focused on in Version 1.0, including the accuracy of the standardized uptake value (SUV), effect of body size together with adjustment of scanning duration,

and time-of-flight (TOF) reconstruction technique. Version 2.0 also presents data acquired with new PET camera models that were not tested in Version 1.0. Reference values for physical indicators of phantom image quality have been updated as well.

**Keywords** Guideline · FDG-PET · Oncology · Noise equivalent count · Phantom

### Objective

The objective of this guideline is to define the criteria for the data acquisition protocol for oncology FDG-PET (PET/CT) scans in order to standardize the PET image quality among PET centers and different PET camera models. It describes the method for phantom experiments and human image quality evaluation and provides recommended values as a reference. The optimum imaging protocol for each camera model can be determined by using this guideline as a manual, and by comparing the results with the recommended values.

The Version 1.0 (*Kakuigaku-Gijutsu* 2009; 29:195–235) and the English synopsis [1] did not deal with the accuracy of SUV values, and did not provide references for scanning patients with large body weight, which inevitably degrades image quality and requires longer scanning duration. Furthermore, new reconstruction techniques such as time-of-flight (TOF) and point-spread-function (PSF) have become available, which affect image spatial resolution and noise in a way different from the conventional OSEM reconstruction technique. To address these issues, the joint task force again worked on the data of phantom and patient scans acquired with PET cameras currently used in Japan, including new PET camera models installed after Version

---

H. Fukukita  
International University of Health and Welfare, Ohtawara, Japan

K. Suzuki  
Dokkyo Medical University Hospital, Tochigi, Japan

K. Matsumoto  
Kyoto College of Medical Science, Nantan, Japan

T. Terauchi · N. Shimada  
National Cancer Center, Tokyo, Japan

H. Daisaki  
Nihon Medi-Physics, Tokyo, Japan

Y. Ikari · M. Senda (✉)  
Division of Molecular Imaging, Institute of Biomedical Research and Innovation, 2-2 Minatojima-minamimachi, Chuo-ku, Kobe 650-0047, Japan  
e-mail: senda@fbri.org

1.0 was published. The outcome was published as Version 2.0 in *Kakuigaku-Gijutsu* 2013; 33:377–420, and this article is the English version.

### Need for the guideline

The quality of FDG-PET images acquired with PET camera (either dedicated PET or PET/CT unit) depends on the camera model, injected activity, scanning duration, and other details of the data acquisition protocol. It also depends on body size, with larger subjects generally showing poorer image quality with the same injected activity per body weight and the same scanning duration. An optimum data acquisition protocol has not necessarily been established for each camera model. On the other hand, radiation safety regulations and operational limitations may prevent injecting sufficient activity and/or scanning for a sufficient duration in the clinical setup.

The diagnostic accuracy reported by a PET center may not be applicable to other centers that use other camera models with different data acquisition protocols, and may therefore provide a different image quality. Unless image quality is universally controlled by standardization of the scanning protocol, the FDG-PET scan will not be validated as a reliable diagnostic tool. Multicenter studies and clinical trials are not possible if the image quality depends on the site. When FDG-PET is used as an endpoint in clinical trials of new anticancer drugs, in which the efficacy of treatment is evaluated based on the disappearance or decline of FDG uptake by lesions, it is essential to assure a certain level of image quality.

There is therefore a growing need for a method to determine a data acquisition protocol that provides adequate image quality of a given camera model, as well as standards of image quality evaluation applicable to human FDG-PET images acquired with any camera model.

### Contents of and instructions concerning the guideline

#### Phantom experiments

Phantom experiment #1 of this guideline allows determination of the minimum scanning duration to detect a 10-mm-diameter hot sphere with 1:4 background activity, simulating a subject of standard size injected with 3.7 (or 7.4) MBq/kg FDG and imaged at 1 h post-injection. In Phantom experiment #2, hot spheres of various sizes with 1:4 background activity are imaged in a given data acquisition protocol for the evaluation of visualization, as well as under noise-free conditions to estimate image resolution based on the recovery coefficient (RC).

The reconstruction condition, which affects image quality and spatial resolution, may be predetermined by the users or the manufacturer, but can also be determined with phantom experiments.

Since detection of a 10-mm hot sphere with 1:4 background activity is a challenging goal, a routine data acquisition protocol may be determined apart from the phantom experiments considering the clinical requirements and operational limitations of the PET center.

Phantom experiments use a body phantom of the standard size (30 cm wide), which does not provide direct evidence for thicker or thinner subjects. While body phantoms of other sizes are not readily available, this guideline Ver 2.0 presents data of specially designed body phantoms simulating larger subjects (33 and 36 cm wide), from which recommendations regarding the required scanning duration for larger subjects could be derived.

#### Human image quality evaluation

The clinical part of the guideline defines physical parameters ( $NEC_{\text{patient}}$ ,  $NEC_{\text{density}}$  and liver SNR) and proposes their recommended values as an easy and objective reference for the image quality of human whole-body FDG-PET. These three parameters are used in this guideline, because they are believed to be good indicators of image quality [2]. These reference values, however, may depend on the PET camera model and the subject body size to some extent. The human image quality is also influenced by various subject factors including blood glucose level, resting conditions and body motion. Therefore, human images should finally be checked visually by a qualified physician or technologist.

#### Coverage of PET scanner types and acquisition modes

Although this guideline is designed for application to a PET/CT scanner in 3D data acquisition mode, which is the norm for oncology scans at present, it can be applied to a dedicated PET scanner as well as to a scanner operated in 2D data acquisition mode. These are collectively referred to as PET camera in this guideline. A scanner with continuous bed movement for a simultaneous emission and transmission scan is also evaluable with this guideline, although the results may require cautious interpretation. As Phantom experiment #1 defined in this guideline requires list mode acquisition, an alternative method is described in Version 1 of this guideline for a scanner that does not provide a list mode acquisition mode. This guideline also requires measurement of prompt and random count rates, for which consultation with the manufacturer may be necessary.

## Phantom experiments procedure and evaluation criteria

This section describes two experiments (#1 and #2) using  $^{18}\text{F}$ -solution and an IEC body phantom (image quality phantom) referred to in the NEMA NU-2 2007 Standard [3]. Another phantom (scatter phantom of the NEMA NU-2 2007 Standard) may be placed adjacent to the body phantom to account for the activity outside the field of view, which is preferable but not essential in this guideline.

If little information is available as in the case of a new scanner model/version or if a new reconstruction parameter is applied, Phantom experiment #1 has to be carried out beforehand to obtain the optimum data acquisition conditions followed by Phantom experiment #2. If a data acquisition protocol is already in use, Phantom experiment #1 may be skipped and image quality can be confirmed by Phantom experiment #2 under that protocol.

### Phantom experiment #1

#### Outline

Since lesion detectability in a PET image and overall image quality depend on the count statistics, Phantom experiment #1 determines an appropriate scanning duration that enables visualization of a 10-mm-diameter hot sphere of unknown localization embedded in a warm background of 1:4 activity concentration ratio. The lid, to which the sphere is attached, is screwed on at an arbitrary angle so that only the person who has prepared the phantom knows the localization of the 10-mm sphere. Data are acquired by list mode, from which PET images of various data acquisition duration (1–10 min) are reconstructed and evaluated for detectability of the hot sphere.

#### Data acquisition

**Phantom preparation** Measure the background volume of the phantom beforehand. Using a regularly checked dose calibrator and taking decay into consideration, prepare FDG with sufficient activity to make a background concentration of 5.3 kBq/ml at the start of data acquisition. Fill exactly one-fourth of the background volume with tap water, add the entire FDG that was precisely measured for the activity and stir to make a hot solution. Draw an aliquot and put it into the 10-mm sphere. If Phantom experiment #2 is to follow, draw another 60 ml of the solution for later use. Fill up the phantom background with tap water and stir to make a warm solution. Fill the other five spheres with the warm background solution.

**Scanning** Place the body phantom horizontally on the bed so that the hot spheres are localized at the center of

the field of view in the  $z$ -axis. Start acquiring two sets of data in list mode, each for 12 min, exactly when the background activity concentration has decayed to 5.30 and 2.65 kBq/ml, respectively. Record prompt and random coincidence counts at the same time. Reconstruct PET images of 1, 2, 3,...10 min data acquisition duration (scanning duration), three sets for each duration, by summing the data starting at 0, 1, 2 min and lasting for 1, 2, 3,...10 min. Use image reconstruction parameters that are routinely used or recommended for the camera model.

#### Evaluation

PET image quality is evaluated for each scanning duration with (1) visual score, (2) phantom noise equivalent count ( $\text{NEC}_{\text{phantom}}$ ), (3) % contrast ( $Q_{\text{H},10\text{mm}}$ ), (4) % background variability ( $N_{10\text{mm}}$ ) and (5) phantom SUV quantitation ( $\text{SUV}_{\text{B,ave}}$ ).  $\text{NEC}_{\text{phantom}}$ ,  $Q_{\text{H},10\text{mm}}$ , and  $N_{10\text{mm}}$  are computed based on the NEMA Standards. Definition and derivation of these physical indicators are described in “Appendix”.

The PET images are visually evaluated regarding the detectability of the 10-mm-diameter hot sphere in a three-step (0, 1, 2) scale by one or more JSNM-certified PET physicians, who do not know the hot sphere localization or the slice on which it is to be visualized. If JSNM-certified PET physicians are not available, JSNM board certified nuclear physicians or JSNMT certified nuclear technologists may assume this role. The images are examined in ascending order of scanning duration on the actual viewer/computer being used clinically. The images are displayed using an inverse gray scale with an upper level of  $\text{SUV} = 4$ , which equals the activity concentration of the hot sphere, and a lower level of  $\text{SUV} = 0$ . All the slices should be looked at. The image is scored 2 if the hot sphere is “identifiable”, 1 if it is “visualized, but similar hot spots are observed elsewhere”, and 0 if it is “not visualized”. The score is averaged across the three image sets for each scanning duration and across the physicians.

#### Recommendations

This guideline recommends the scanning duration that provides an image with an average score of 1.5 or more, i.e., the 10-mm hot sphere is detected in/by half or more of the cases/readers. The physical indicators may be used as a reference when determining the optimum scanning duration; the reference values are  $\text{NEC}_{\text{phantom}} > 10.8$  (Mcounts),  $N_{10\text{mm}} < 5.6$  (%), and  $Q_{\text{H},10\text{mm}}/N_{10\text{mm}} > 2.8$  (%), and  $\text{SUV}_{\text{B,ave}}$  should be close to unity (theoretical value). Supporting data for these reference values are presented in Sect. “Phantom experiment #1”.

## Phantom experiment #2

### Outline

In Phantom experiment #2, a body phantom containing hot spheres of various sizes is imaged with a given clinical data acquisition protocol to evaluate their visualization as well as to evaluate the image uniformity in the background area. The phantom is also imaged in a noise-free condition to estimate image resolution based on the recovery coefficient (RC) of the spheres. Phantom experiment #2 can either be carried out following #1 or separately. In the former case, the scanning duration should be adjusted to account for radioactivity decay.

### Data acquisition

**Phantom preparation** A body phantom is prepared in the same way as in Phantom experiment #1, except that all six (10-, 13-, 17-, 22-, 28- and 37-mm diameter) hot spheres are filled with hot solution. The background is filled with 1:4 warm activity concentration like in Phantom experiment #1.

**Scanning** The phantom is scanned twice; namely, in the given clinical condition and in a noise-free condition.

In the given clinical condition, the scanning duration is determined so that equivalent counts are obtained assuming that the phantom simulates a 60 kg subject injected with 222 MBq (3.7 MBq/kg) FDG. If a 60-kg subject is injected more (or less) activity than 222 MBq in the given protocol, the scanning duration is accordingly shortened (or elongated) inverse-proportionally. The scan starts when the activity concentration decays to the following value. If experiment #2 is done alone, the emission scan starts when the activity concentration decays to 2.65 kBq/ml (within  $\pm 5\%$ ). If experiment #2 is done following #1, the emission scan starts when the activity concentration decays to 1.325 kBq/ml (within  $\pm 5\%$ ), taking twice the scanning duration. When setting up the scan, input the phantom volume as “patient weight (kg)”, and the activity at the start of scan as “injected activity”.

After the static scan of the given clinical condition, a second scan of 30 min duration is carried out as a noise-free condition to measure the recovery coefficient.

With all those scans, an acquisition method should be selected that enables the recording of prompt and random coincidence counts in a readable format in the sinogram header or in a separate file. The image reconstruction parameters used in the usual clinical diagnostic scans should be applied to the phantom experiments.

### Evaluation

The quality of PET image acquired in the clinical condition is evaluated by (1) visual inspection regarding visualization

of each sphere, (2) phantom noise equivalent count ( $NEC_{\text{phantom}}$ ), and (3) % contrast ( $Q_{H,10\text{mm}}$ ) and (4) % background variability ( $N_{10\text{mm}}$ ) for the 10-mm-diameter sphere.

The recovery coefficient for a j-mm-diameter hot sphere ( $RC_j$ ) is calculated as the maximum pixel value ( $C_j$ ) within the region of interest (ROI) over the sphere on the reconstructed image acquired in a noise-free condition divided by that of the 37-mm-diameter sphere:  $RC_j = C_j/C_{37}$ .

### Recommendations

Images acquired under clinical conditions should preferably provide visualization of the 10-mm-diameter sphere and the physical indicators of  $NEC_{\text{phantom}} > 10.8$  (Mcounts),  $N_{10\text{mm}} < 5.6$  (%), and  $Q_{H,10\text{mm}}/N_{10\text{mm}} > 2.8$ , which are the same criteria as in Phantom experiment #1 in Sect. “[Recommendations](#)”, and  $SUV_{B,\text{ave}}$  should be close to unity (theoretical value).

A reconstruction condition that provides a spatial resolution of 10 mm FWHM or better ( $RC_{10\text{mm}} > 0.38$ ) is recommended (see Sect. “[Simulation of image resolution and Phantom experiment #2](#)”).

## Evaluation of human PET image quality

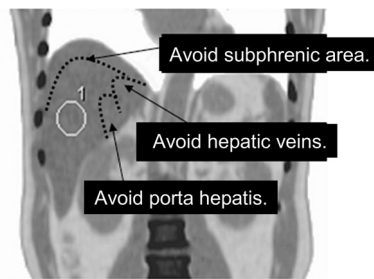
### Objective

This section describes the clinical part of the guideline, in which physical indicators of image quality of human whole-body FDG-PET are defined, including  $NEC_{\text{patient}}$  (noise equivalent count per axial length),  $NEC_{\text{density}}$  (NEC per volume) and liver SNR (mean/SD within liver ROI), together with their reference values as recommended criteria.

While it is preferable that human images be acquired under conditions that meet the recommended criteria of Phantom experiment #2, especially that for image resolution ( $RC_{10\text{mm}} > 0.38$ ), this guideline recommends criteria for the physical parameters that are directly measurable on human data, considering the inherent limitations of the phantom experiments such as body size variations.

### Method

The criteria are applicable to whole-body FDG-PET images covering the area from at least the neck to the abdomen. The images should be acquired while recording the prompt and random coincidence counts in each bed position. The transmission or X-ray CT images should also be generated together with PET images to compute length and volume.



**Fig. 1** How to place ROI over the liver

For the whole-body image, bed positions corresponding to the axial span from the neck to the abdomen are determined by excluding the brain and urinary bladder. Prompt and random counts are extracted for each bed position, from which  $NEC_{\text{patient}}$  and  $NEC_{\text{density}}$  are computed (see [Appendix](#)). The liver SNR is computed as mean/SD within the liver ROI that is placed separate from the porta hepatis and major vessels in three coronary sections (Fig. 1).

### Recommendations

This guideline recommends that the physical indicators meet the criteria of  $NEC_{\text{patient}} > 13$  Mcounts/m,  $NEC_{\text{density}} > 0.2$  kcounts/cm<sup>3</sup> and liver SNR > 10.

Since these reference values may strictly depend on the camera model, they may be subject to future modifications and revisions. It may also be inappropriate to use the criteria if FDG distribution is far from normal, such as in cases in which lesions show extremely strong FDG accumulation.

## Discussion

### Dependence on camera model

This guideline aims to establish standards of indicators to assure image quality independent of the camera model. Standards of  $N_{10\text{mm}} < 5.6$  (%),  $Q_{H,10\text{mm}}/N_{10\text{mm}} > 2.8$  and  $NEC_{\text{phantom}} > 10.8$  (Mcounts) have been proposed for the phantom image quality parameters based on the results of Phantom experiment #1 for a number of camera models regarding detection of a 10-mm hot sphere of unknown localization with 1:4 background activity. These reference values have been revised in the present Version 2.0 based on the data pertaining to recent camera models. The image spatial resolution should be better than 10 mm FWHM corresponding to  $RC > 0.38$  for the 10-mm sphere in Phantom experiment #2. As for human images, image quality parameters of  $NEC_{\text{patient}} > 13$  Mcounts/m,  $NEC_{\text{density}} > 0.2$  kcounts/cm<sup>3</sup> and liver SNR > 10 have tentatively been proposed as the minimum standards based

on the clinical data at a number of PET centers. Although these standards may depend on the camera model, our results suggest that they may be roughly applicable to all camera models.

Computation of NEC needs scatter fraction, which was obtained from the literature or measured under the conditions defined in the NEMA standard and was not measured concurrently in each phantom experiment or human scan in the present study. Therefore, the scatter fraction value may have an error, which may be one of the reasons for the camera dependence of the relationship between NEC and visual score.

### Scatter fraction

The scatter fraction of a PET camera depends on factors such as the camera model, acquisition mode, body size, and activity outside the field of view [4]. In general, the scatter fraction measured with a scatter phantom based on NEMA standard may provide a lower value than clinical scans, because it increases as the subject size increases [5]. Moreover, the scatter fraction is related to the energy lower level discriminator (ELLD) and is reported to be higher than 40 % if ELLD is set below 400 keV [6, 7]. In addition, the scatter fraction is influenced by the radioactivity concentration if the PET camera detector contains lutetium (<sup>176</sup>Lu), and the data are acquired in 3D mode [8]. Therefore, the scatter fraction varies widely with body size and activity inside or outside the direct field of view. However, since the real-time measurement of the scatter fraction is impossible with clinical scans, this guideline instructs using the scatter fraction values based on NEMA NU 2-2007 as an intrinsic value for each camera model. Therefore, there is a possibility of errors in the actual scatter fraction for each human scan.

### Relationship between phantom results and human scanning conditions

In many PET centers in Japan, patients are injected with 3.7 MBq/kg FDG and are scanned starting 60 min post-injection. Suppose that the target region is scanned at 68 min post-injection (physical decay to 65 %). Assuming that 20 % of injected FDG is excreted in the urine [9], and that the remaining FDG is distributed uniformly within the body except the adipose tissue, which constitutes 27 % of the total body volume [10], the soft tissue activity concentration is estimated to be  $3.7 \text{ MBq/kg} \times 1 \text{ kg/l} \times 0.65 \times 0.8/0.73 = 2.64 \text{ MBq/l}$ , which is comparable to the background activity concentration in the phantom experiment (2.65 kBq/ml) (specific gravity = 1). The soft tissue SUV value is then  $0.8/0.73 = 1.1$ , which is compatible with the SUV value in the mediastinum or abdomen



observed in routine clinical experience. The cross-sectional area of the body phantom (550 cm<sup>2</sup>) corresponds to that of a standard Japanese with a body weight of 60 kg. Therefore, the body phantom at an activity concentration of 2.65 kBq/ml corresponds to a standard Japanese subject of 60 kg injected with 3.7 MBq/kg FDG and scanned starting 60 min post-injection, and Phantom experiment #1 corresponds to determining the minimum scanning duration to detect a 10-mm hot lesion with 4 times the background activity concentration in such a subject.

The results of the present study indicated that scanning for 3–4 min or longer is necessary for many camera models, except for recent ones, to visualize a 10-mm sphere in Phantom experiment #1. This is longer than 2–3 min, which is usually adopted for a standard sized subject in Japan. This suggests that a 10-mm lesion with 1:4 background activity may not be visualized in routine clinical scans except for some recent PET camera models. As a matter of fact, considering that the image activity of a 10-mm hot lesion with 4 times the background is decreased to  $SUV = 1.7$  by the partial volume effect, it may not be easy to detect a 10-mm lesion of  $SUV = 1.7$  of unknown localization on PET images alone in a routine clinical situation.

#### Body size and current data acquisition protocol

More activity was injected in heavier subjects in the routine clinical setup of all the PET centers surveyed by the task force, and some centers further increased the scanning duration in subjects with high body weight or BMI [=weight (kg)/height (m)/height (m)]. The results of the present study indicated a trend of image quality degradation as the body weight or BMI increased as long as non-TOF reconstruction was employed, suggesting that in general, the current routine protocol adjustment for increased body size may not be sufficient when using conventional OSEM reconstruction algorithm. It is advisable to inject more activity or (because injecting more activity may not work due to increased random rate) to increase scanning duration in large-size subjects to acquire equivalent image quality as in small-size subjects. Readjustment of acquisition duration may be advisable especially for subjects with 25 or larger BMI because some cases in this category were found to present lower  $NEC_{\text{patient}}$  and  $NEC_{\text{density}}$  than the recommended values. Interestingly, visual image quality of routine clinical scans did not depend on BMI among images reconstructed with TOF algorithm, which suggests that TOF is effective in reducing noise and improving image quality for large-size patients even with the same count statistic (see Sect. “Human image quality evaluation”).

Body phantoms of larger size were designed by the task force to examine the effect of object size directly on the

visualization of spheres and on the physical indicators of image quality. The data have indicated that scanning time should be elongated to obtain equivalent image quality in larger phantoms containing the same activity concentration as compared with the phantom of standard size. As tentative estimates of recommendation based on the relationship between body weight and cross section in the Japanese population, 1.3, 1.7 and 2.2 times as many counts (long scanning duration) are necessary for patients of 70, 80, and 90 kg body weight, respectively, as compared with 60 kg patients, if injected with the same activity per body weight (see Sect. “Effect of phantom size on image quality”).

#### Supporting data

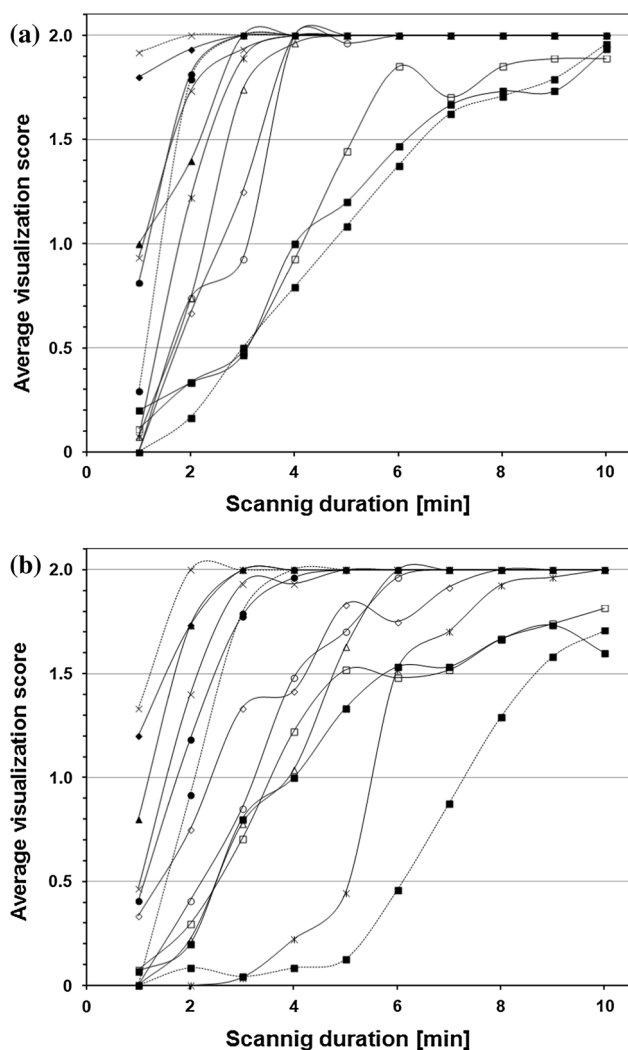
This section presents phantom and human data on a number of PET camera models acquired and/or evaluated based on this guideline, from which the recommended reference values have been derived.

#### Phantom experiment #1

**Methods** Phantom experiment #1 was carried out according to this guideline on 13 PET camera models (Aquiduo, Biograph LSO, Discovery ST, Discovery STE, Discovery STEP, SET-3000 B/L, SET-3000 G/X, Biograph mCT, Discovery 600, Discovery 690, GEMINI GXL, GEMINI TF, SET3000 GCT/M) to determine the optimum scanning duration and to investigate the validity of the physical parameters as indicators of the 10-mm hot sphere visualization. The reconstruction condition, which is routinely used in the PET center that housed the PET camera, was employed for this experiment. The PET images were visually evaluated by nine physicians and technologists using “Fusion Viewer 2.0” (Nihon Medi-Physics) software to derive visualization scores.

**Results and discussion** Figure 2 represents the relationship between the average score of visualization for the 10-mm-diameter hot sphere and the scanning duration. As the scanning duration increased, the visualization of each PET camera model improved, although the optimum duration depended on the model.

Figure 3 represents the relationship between the average score of visualization for the 10-mm-diameter hot sphere and the physical parameters. The  $NEC_{\text{phantom}}$ ,  $N_{10\text{mm}}$  and  $Q_{H,10\text{mm}}/N_{10\text{mm}}$  were similarly related to the visual score regardless of the camera model, suggesting the validity of those parameters as indicators of the hot sphere detectability. As scanning duration increased,  $NEC_{\text{phantom}}$  increased and  $N_{10\text{mm}}$  decreased, with both contributing to improving the image quality and lesion detectability. On the other hand,  $Q_{H,10\text{mm}}$  was poorly associated with the visual



**Fig. 2** Relationship between scanning duration and visualization score for 10-mm sphere in Phantom experiment #1 (a 5.30 kBq/ml, b 2.65 kBq/ml). Symbols represent camera models

score, as it approached a constant when a certain level of counts were acquired. It should be noted that  $N_{10\text{mm}}$  and  $Q_{H,10\text{mm}}$  are affected by the reconstruction condition while  $NEC_{\text{phantom}}$  is not, and that the reconstruction condition was predetermined in the present experiments. Therefore, different results may have been obtained under different reconstruction parameters even with the same PET camera model.

A few PET camera models presented very poor detectability of the 10-mm sphere, possibly due to large  $N_{10\text{mm}}$ , i.e., poor image uniformity in the background area. This may be improved by installing software that enhances corrections for detector efficiency normalization and for attenuation and scatter. Suppression of  $N_{10\text{mm}}$  is especially important because detection of a hot sphere requires perception of the hot sphere activity in contrast to the surrounding false positive noise activities.

In this guideline Version 2.0, new PET cameras having TOF reconstruction algorithm were also examined.

The median value of the 13 camera models that provided the average visual score of 1.5 in this experiment was adopted as the recommended reference value for each of the three physical indicators:  $NEC_{\text{phantom}} > 10.8$  (8.7–17.5) and  $> 8.8$  (6.9–13.2) (Mcounts),  $N_{10\text{mm}} < 5.6$  (4.2–10.6) and  $< 6.3$  (5.8–8.1)(%), and  $Q_{H,10\text{mm}}/N_{10\text{mm}} > 2.8$  (2.1–3.2) and  $> 2.2$  (2.1–2.8), for 5.30 and 2.65 kBq/ml concentration, respectively (95 % confidence interval in parenthesis).

#### Simulation of image resolution and Phantom experiment #2

Computer simulation was carried out to determine the relationship between spatial resolution and the recovery coefficient measured under noise-free conditions in Phantom experiment #2. Using a 3D Gaussian filter with FWHM = 10 mm, the recovery coefficients of the spheres under the present experimental conditions turned out to be:  $RC_{10\text{mm}} = 0.38$ ,  $RC_{13\text{mm}} = 0.52$ ,  $RC_{17\text{mm}} = 0.72$ ,  $RC_{22\text{mm}} = 0.88$  and  $RC_{28\text{mm}} = 0.97$  (Fig. 4). Based on this simulation,  $RC_{10\text{mm}} > 0.38$  was adopted as the recommended reference value in this guideline, assuming that a spatial resolution of 10 mm FWHM or better would be necessary for an oncology FDG-PET image with sufficient quality.

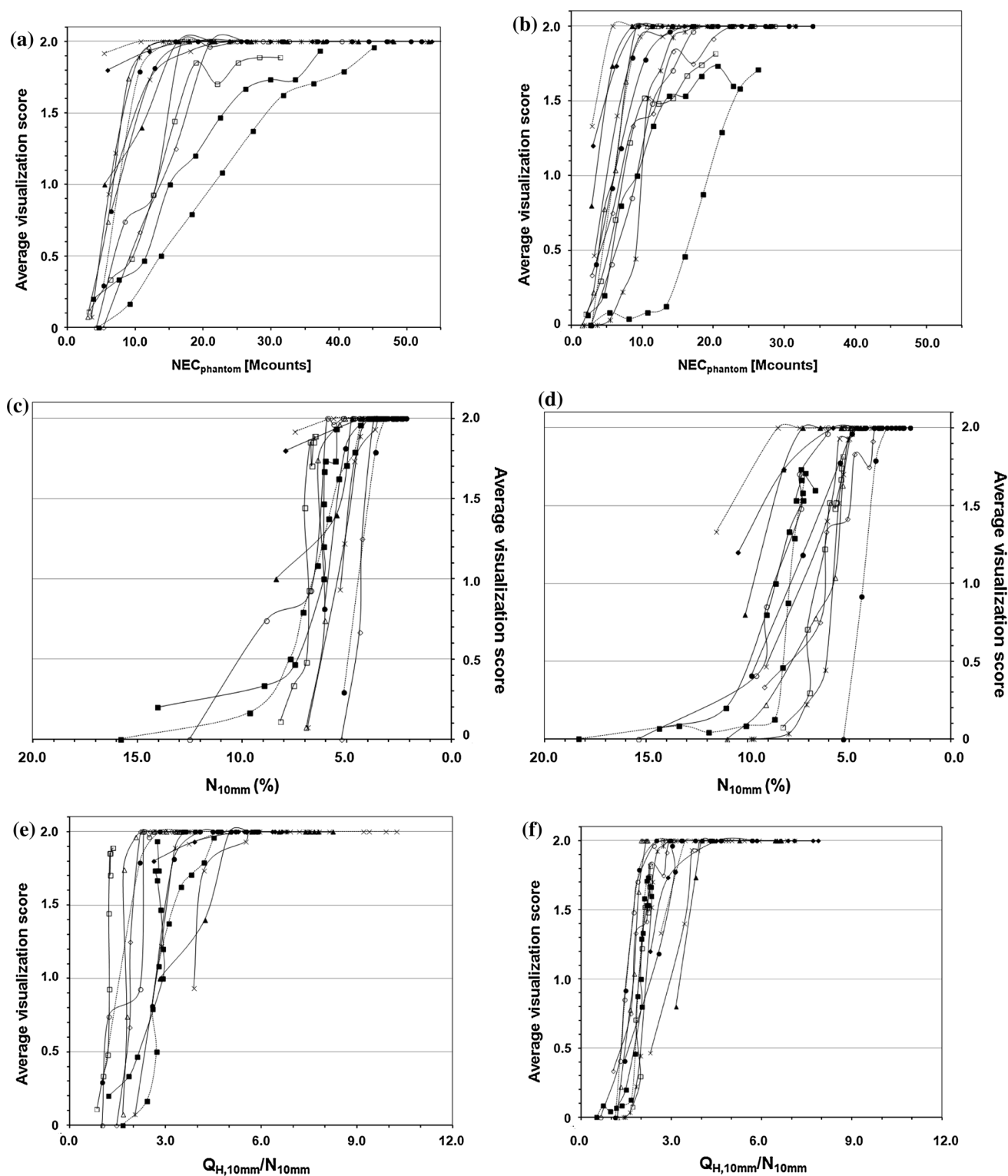
All the PET camera models examined in this study met the requirement by selecting appropriate reconstruction parameters.

The so-called Gibbs Ringing artifact was frequently observed in the images reconstructed with PSF algorithm, i.e., RC values larger than 1.0 were observed for 17 and/or 22 mm spheres (object size being 3–4 times as large as the crystal size). PSF reconstruction should be treated with caution in quantitative measurement although it is considered to improve lesion detectability by emphasizing edges.

#### Accuracy of phantom background SUV ( $SUV_{B,ave}$ )

**Methods** Quantitative capability was examined on the 13 PET camera models mentioned in Sect. “Phantom experiment #1” using a standard body phantom prepared in the procedure of Phantom experiment #1. The phantom was scanned for 10 min starting at the concentration of 5.30 kBq/ml without a scatter phantom. Images were reconstructed with the usual parameters and average SUV in the background area was obtained using Fusion Viewer (Ver.2.0).

**Results** Table 1 presents  $SUV_{B,ave}$  of each camera model. The median of the 13 PET camera models was 1.01 (95 % confidence interval 0.98–1.05).



**Fig. 3** Relationship between visualization score for 10-mm sphere and  $NEC_{\text{phantom}}$  (a, b),  $N_{10\text{mm}}$  (c, d), and  $Q_{H,10\text{mm}}/N_{10\text{mm}}$  (e, f) in Phantom experiment #1 for activity concentration of 5.30 kBq/ml (a, c, e) and 2.65 kBq/ml (b, d, f). Symbols represent camera models

$SUV_{B,\text{ave}}$  depends on the details of image reconstruction and other correction procedures, injected activity and cross-calibration. The accuracy of  $SUV_{B,\text{ave}}$  is influenced

by the frequency of cross-calibration, clock synchronization, and other equipment maintenance and facility management. In this study, we also examined the data





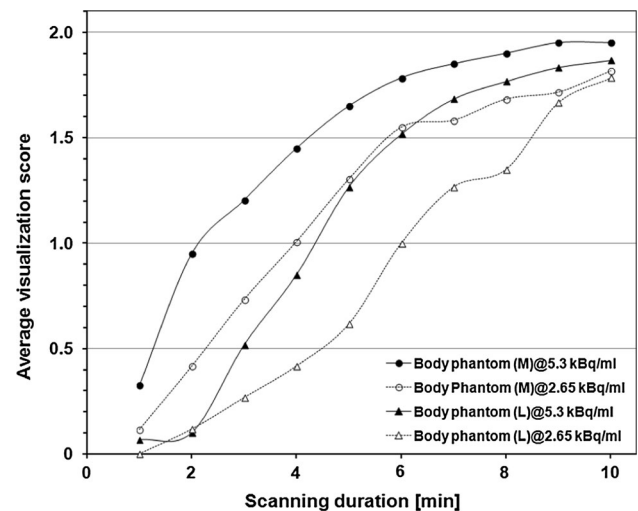
**Fig. 4** Simulated image of digital body phantom generated with a Gaussian filter of 10 mm FWHM isotropic image resolution

**Table 1** Background SUV ( $SUV_{B,ave}$ ) for each camera model (theoretical value = 1.00)

Camera model	$SUV_{B,ave}$
Aquiduo	1.03
Biograph LSO	0.99
Biograph mCT	1.03
Discovery ST (3D)	0.97
Discovery STE	1.02
Discovery STEP	1.01
Discovery 600	0.95
Discovery 690	1.04
GEMINI GXL	1.00
GEMINI TF	0.98
SET-3000B/L	1.13
SET-3000GCT/M	0.95
SET-3000G/X	1.13

acquisition methods, state of the PET camera maintenance and software updating. The results indicated that  $SUV_{B,ave}$  obtained by SET3000 B/L and SET3000 G/X was more than 10 % off the theoretical value (1.0), of which visual inspection of the images also showed non-uniform activity in the center of background area. This may be caused by inappropriate correction methods for normalization, attenuation and scatter, and  $SUV_{B,ave}$  and uniformity may be improved by installation of new software which has recently been released. As a matter of fact,  $SUV_{B,ave}$  was improved in GEMINI TF after upgrading SUV calibration software.

$SUV_{B,ave}$  allows evaluation of the overall accuracy of the quantitative capability based on the PET camera performance, data acquisition quality, and maintenance and management of the facility. Reliability of lesion SUV values may be evaluated with  $SUV_{B,ave}$  and RC for the sphere of corresponding size measured in Phantom experiment #2. In clinical settings, however, not only partial volume effect and quantitative capability but also physiological factors such as motion and respiration affect SUV



**Fig. 5** Relationship between scanning duration and visualization score for 10-mm sphere in Phantom experiment #1 on larger (M: 33 cm and L: 36 cm) phantoms containing activity concentration of 5.30 and 2.65 kBq/ml. Average of four camera models

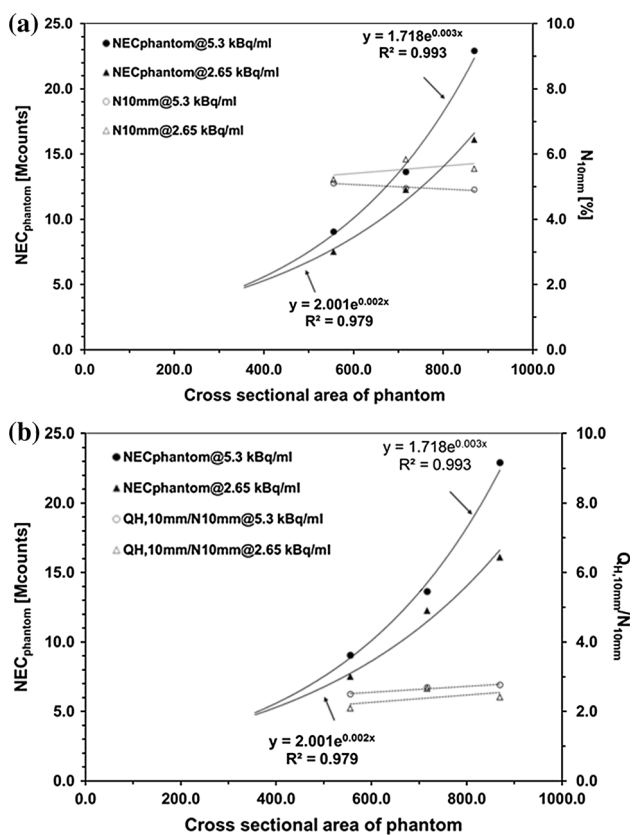
values, and accurate measurement of lesion SUV is nearly impossible.

In summary, measurement of SUV using a phantom containing an area with theoretical SUV of 1.0 allows evaluation of quantitative capability of the PET camera together with the calibration system. If it is off 1.0, causes should be investigated and corrective measures taken.

#### Effect of phantom size on image quality

**Methods** Four PET camera models (Discovery STEP, Discovery 600, Biograph LSO, Aquiduo) were tested for the effect of object size on the image quality using two additional larger body phantoms (33 and 36 cm in major axis, corresponding to body weight of 80 and 100 kg, respectively) that were designed to be similar to the standard NEMA IEC body phantom (30 cm in major axis corresponding to 60 kg) except that the background pool activity area is larger. Data were acquired according to Phantom experiment #1 (5.3 and 2.65 kBq/ml) without a scatter phantom, and 60 sets of images were obtained with the usual reconstruction parameters. The images were evaluated by five readers with Fusion Viewer Ver. 2.5. ROI was placed and physical parameters were computed using “PETquant” (Ver 2.02.02). Phantom experiment #2 was also carried out and RCs were computed. The results were averaged across the four PET cameras.

**Results** Figure 5 illustrates the relationship between scanning duration and visualization score of the 10-mm sphere in Phantom experiment #1 on the larger phantoms. Longer scanning duration was required to provide



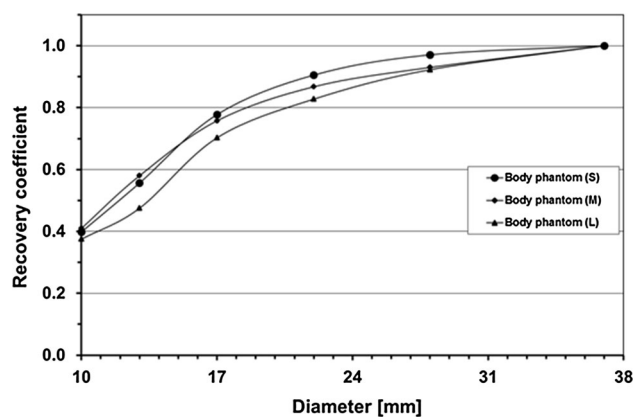
**Fig. 6** Relationship of phantom cross-sectional area against reference values of  $NEC_{phantom}$ ,  $N_{10mm}$  (a), and  $Q_{H,10mm}/N_{10mm}$  (b) that make visualization score  $>1.5$  in Phantom experiment #1 on standard and larger phantoms containing activity concentration of 5.30 and 2.65 kBq/ml. Average of four camera models

detectability of the 10-mm sphere in larger phantoms containing the same radioactivity concentration.

Figure 6 presents the relationship of the cross-sectional area of the phantom against reference values of  $NEC_{phantom}$ ,  $N_{10mm}$ , and  $Q_{H,10mm}/N_{10mm}$  that made the visualization score  $>1.5$ . As the phantom became larger, larger  $NEC_{phantom}$  was required while similar  $N_{10mm}$  and  $Q_{H,10mm}/N_{10mm}$  were sufficient to visualize the 10-mm sphere. This suggests that  $N_{10mm}$  and  $Q_{H,10mm}/N_{10mm}$  remain good indicators of lesion detectability irrespective of the object size. Although NEC is believed to reflect image quality, NEC density, which is used for evaluation of patient scans, may be a better indicator for variable object size.

Figure 7 illustrates RC curve against the sphere diameter measured in Phantom experiment #2, which showed a tendency of lower RC for larger phantoms.

Image noise inevitably increases due to increased scatter and attenuation in larger phantoms, which hinders detection of hot spheres as contrasted to the background noise, and thus the results of the present study. The large phantom used here contained the same radioactivity concentration as



**Fig. 7** Recovery coefficients (RCs) for hot spheres of various diameters obtained in noise-free scans in Phantom experiment #2 on standard (S: 30 cm) and larger (M: 33 cm, L: 36 cm) phantoms. Average of four camera models

the standard phantom, corresponding to the same injected activity per body weight in patient scans. Therefore, longer scanning duration should be necessary to obtain the same lesion detectability in patient scans if injected with the same activity per body weight. Increasing the injected activity per body weight is an alternative if radiation exposure permits and the dose is available, although it may not increase NEC and image quality as much as elongation of scanning duration does because of increased random rate and count loss. TOF reconstruction may be another solution as described in Sect. “Human image quality evaluation”.

Based on the relationship between body weight and cross-sectional area of the Japanese population, the results of the present study allow estimation of the necessary scanning duration to obtain the same lesion detectability in patients of large body weight injected with the same activity per body weight: compared with 60 kg body weight as reference, 1.3, 1.7 and 2.2 times as long scanning duration (i.e., many NEC) is required for 70, 80 and 90 kg body weight, respectively.

#### Human image quality evaluation

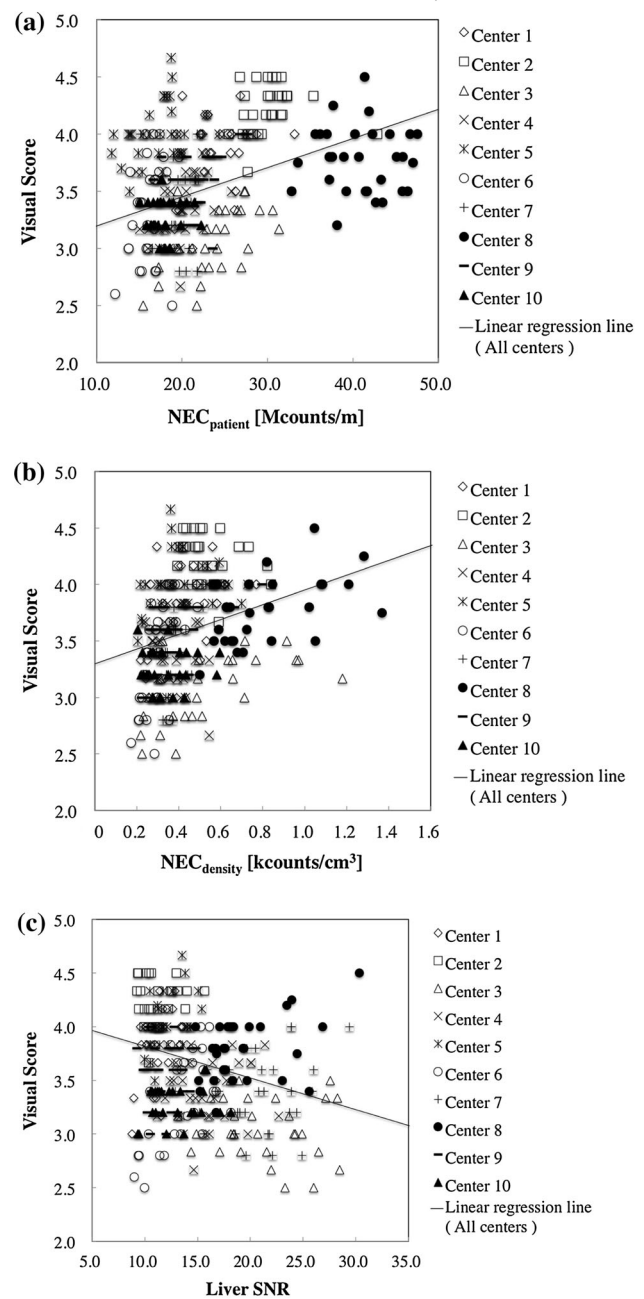
**Methods** To examine the image quality of whole-body FDG-PET images currently acquired clinically in Japan and the relationship with the physical parameters, patient images were collected from 10 PET centers using 10 different PET camera models, 28–30 cases from each center. Those images had been acquired as routine diagnostic scans according to the protocol of each PET center without any artifacts or other problems, and interpreted by local PET physicians and reported to the attending physicians. Images with extremely abnormal FDG accumulation were excluded.

The quality of the images was visually evaluated by five JSNM-certified PET physicians using 5-step scores regarding how and whether they had sufficient quality to be read and interpreted. The image was given a score of 5 for “very good quality”, 4 for “sufficiently good quality”, 3 for “scarcely sufficient quality”, 2 for “not sufficient quality”, and 1 for “unreadable”.  $NEC_{patient}$ ,  $NEC_{density}$  and liver SNR were computed as described above and were compared with the visual score as well as the BMI of the patient. The results were also analyzed separately for PET cameras using time-of-flight (TOF) reconstruction and for non-TOF reconstruction.

**Results and discussion** Figure 8 illustrates the plots of the average visual score against  $NEC_{patient}$ ,  $NEC_{density}$  and liver SNR. These three indicators are known to be excellent indicators of image quality for whole-body FDG-PET/CT images acquired with a single PET camera model [2]. In the present study, in which image data by 10 PET camera models were merged, visual score presented a weak but significant correlation with  $NEC_{patient}$  ( $r = 0.376$ ,  $p < 0.001$ ) and with  $NEC_{density}$  ( $r = 0.432$ ,  $p < 0.001$ ). This suggests that  $NEC_{patient}$  and  $NEC_{density}$  may still be useful indicators of image quality even across different PET cameras and PET centers.  $NEC_{density}$  is less influenced by the body size and by the arm position, which might have provided higher correlation coefficients. On the other hand, liver SNR was weakly and negatively correlated with the visual score ( $r = -0.278$ ,  $p < 0.001$ ). Because liver SNR depends on the image reconstruction method and parameters, which is variable among PET camera vendors and models but is already optimized to some extent, an opposite correlation might have been observed. It is also noted that liver SNR relies on careful ROI placement.

Since the images were all selected from routine clinical scans, heavier patients had been injected with more activity and/or were scanned for a longer duration, so that they would not include images with too high or too low quality. This may be another reason for the weak correlation between the visual score and the physical parameters.

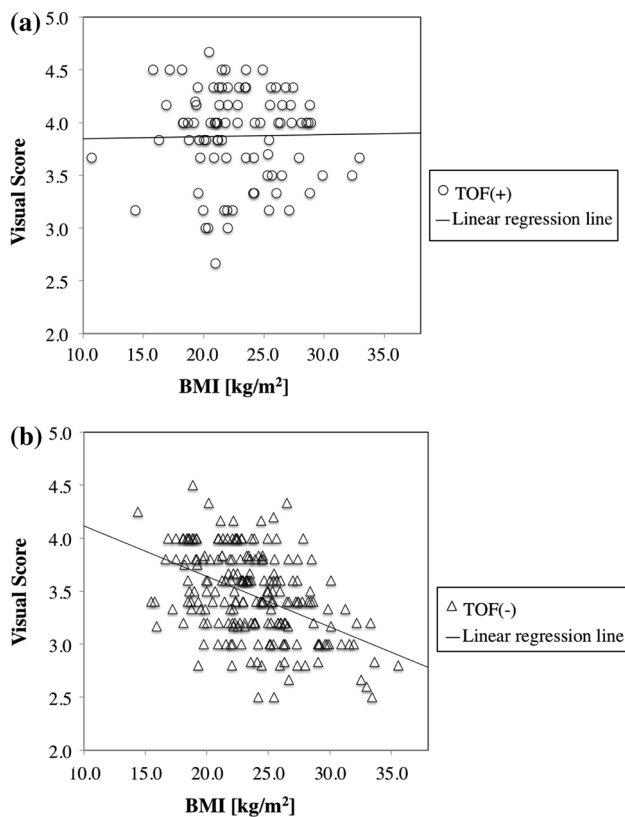
There was a significant difference in the visual score between PET cameras employing TOF reconstruction ( $3.87 \pm 0.44$ ) and non-TOF reconstruction ( $3.46 \pm 0.41$ ), TOF gaining a significantly higher score than non-TOF ( $p < 0.001$ ). No significant difference was observed between TOF and non-TOF for  $NEC_{patient}$  ( $22.4 \pm 6.36$  vs.  $23.5 \pm 8.16$  Mcounts/m) or for  $NEC_{density}$  ( $0.44 \pm 0.13$  vs.  $0.45 \pm 0.22$  kcounts/cm<sup>3</sup>), respectively, which is reasonable because NEC is defined in the acquired raw data and is independent of the reconstruction technique. This also supports the hypothesis that TOF is an effective reconstruction technique for improving image quality of given raw data. Interestingly, liver SNR was lower in TOF than in non-TOF ( $13.1 \pm 2.92$  vs.  $16.0 \pm 5.01$ ,  $p < 0.001$ ),



**Fig. 8** Scatter plots of visual score against  $NEC_{patient}$  (a),  $NEC_{density}$  (b) and liver SNR (c) in patient scans acquired in 10 PET centers. Each plot represents a subject, and *symbols* represent PET centers. Linear regression line for merged data is shown for each graph

suggesting that TOF images may provide higher visual quality even with lower liver SNR.

Figure 9a, b plot visual scores against BMI in TOF images and in non-TOF images, respectively. No correlation was observed for TOF images ( $r = -0.095$ ,  $p = 0.171$ ), while a significant negative correlation was observed for non-TOF ( $r = -0.474$ ,  $p < 0.001$ ). A trend to a lower visual score in patients with larger BMI was pointed out for routine whole-body FDG non-TOF PET/CT scans in our



**Fig. 9** Scatter plots of visual score against body mass index (BMI) for patient scans in 10 PET centers using time-of-flight (TOF) (a) and non-TOF (b) image reconstruction techniques. Linear regression line is shown for each graph

previous study [1], suggesting insufficient adjustment of scanning duration and/or injected activity for large BMI patients, which was also confirmed by the present result in Fig. 9b. However, the lack of this trend for TOF images in Fig. 9a indicates that equivalent visual image quality is obtained for large BMI patients in routine scans and may suggest the effectiveness of TOF in preventing degradation of visual image quality in large BMI patients.

Based on these patient data, the recommended reference values were determined as  $NEC_{\text{patient}} > 13$  (Mcounts/m),  $NEC_{\text{density}} > 0.2$  (kcounts/cm<sup>3</sup>) and liver SNR > 10 for this guideline. It should be noted, however, that these reference values may still depend on the camera model, and that further modification and revision may be necessary to make them reliable criteria for quality control.

### Appendix: physical indicators of image quality

#### Indicators of phantom image quality

In this guideline,  $NEC_{\text{phantom}}$  (noise equivalent count for phantom), percent contrast ( $Q_{H,10\text{mm}}$ ) and percent

background variability ( $N_{10\text{mm}}$ ) are used as indicators of body phantom image quality.

The  $NEC_{\text{phantom}}$  is calculated using the formula in Eq. 1:

$$\begin{aligned} NEC[\text{Mcounts}] &= \frac{T^2}{T + S + (1+k)fR} \\ &= (1 - SF)^2 \frac{(P - D)^2}{(P - D) + (1+k)fD} \\ f &= \frac{S_a}{\pi r^2} \end{aligned} \quad (1)$$

where  $T$ ,  $S$ , and  $R$  represent true, scatter, and random coincidences acquired within the scanning period, and  $P$  and  $D$  represent prompt and delayed coincidences.  $SF$ ,  $k$ , and  $f$  represent scatter fraction, random scaling factor, and ratio of object size to sonogram, respectively.  $S_a$  and  $r$  represent the cross-sectional area of the phantom and the radius of the detector ring diameter, respectively.

The phantom image is reconstructed with all available corrections applied, using the standard reconstruction algorithm and usual parameters for whole-body studies.

A transverse image centered on the hot sphere(s) is used in the analysis. A circular region of interest (ROI) with a 10-mm diameter is drawn on the 10 mm hot sphere. The ROI analysis tool should take partial pixels into account and also permit movement of the ROI in increments of 1 mm or smaller.

Twelve ROIs of the same size are drawn throughout over the background at a distance of 15 mm from the edge of the phantom, but not closer than 15 mm to any sphere. The ROIs are also drawn on the slices as close as possible to  $\pm 1$  and  $\pm 2$  cm on either side of the central slice, resulting in a total of 60 background ROIs, twelve on each of the five slices. The locations of all ROIs should be fixed between successive measurements. The measured activity in each background ROI is recorded. The percent contrast for the 10 mm hot sphere ( $Q_{H,10\text{mm}}$ ) is calculated as follows:

$$Q_{H,10\text{mm}} = \frac{C_{H,10\text{mm}}/C_{B,10\text{mm}} - 1}{a_H/a_B - 1} \times 100\% \quad (2)$$

where  $C_{H,10\text{mm}}$  and  $C_{B,10\text{mm}}$  are the average measured activity in the ROI for the 10-mm sphere and the average measured activity in all the background 10-mm diameter ROIs, respectively.  $a_H/a_B$  is the activity concentration ratio for the hot sphere to the background.

The percent background variability  $N_{10\text{mm}}$  for the 10-mm sphere is calculated as follows:

$$N_{10\text{mm}} = \frac{SD_{10\text{mm}}}{C_{B,10\text{mm}}} \times 100\% \quad (3)$$

where  $SD_{10\text{mm}}$  is the standard deviation of the background ROI counts for the 10-mm sphere, calculated as follows:



$$SD_{10\text{mm}} = \sqrt{\frac{\sum_{k=1}^K (C_{b,10\text{mm},k} - C_{B,10\text{mm}})^2}{K-1}}, K = 60 \quad (4)$$

Indicators of human image quality

$NEC_{\text{patient}}$  (noise equivalent count per patient height) and  $NEC_{\text{density}}$  (noise equivalent count per volume) are evaluated as potential physical indicators of image quality.

The  $NEC_{\text{patient}}$  is defined to allow for patient height normalization. In this guideline, since the axial scanning range is variable,  $NEC_{\text{patient}}$  is defined as shown in Eq. 5.

$$NEC_{\text{patient}}[\text{Mcounts/m}] = \frac{\sum_{i=1}^I NEC_i}{x/100} \quad (5)$$

where  $NEC_i$  and  $x$  represent NEC for each bed position ( $i$ ) and the length [cm] of the axial field of view to be evaluated ( $i = 1$  to  $I$ ), which extends from the neck to the abdomen in this guideline, respectively.

$NEC_i$  is calculated using the formula in Eq. 6.

$$NEC_i[\text{Mcounts}] = (1 - SF)^2 \frac{(P_i - D_i)^2}{(P_i - D_i) + (1 + k)D_i} \quad (6)$$

where  $P_i$  and  $D_i$  represent prompt and delayed coincidences for each bed position. SF represents scatter fraction measured within the NEMA NU 2-2001 Standard [11], and  $k$  is set to 0 or 1 depending on whether you use variance reduction techniques for estimating a smooth random distribution or use direct random subtraction.

$NEC_{\text{density}}$  is defined as shown in Eq. 7.

$$NEC_{\text{density}}[\text{kcounts/cm}^3] = \frac{\sum_{i=1}^I NEC_i}{V} \quad (7)$$

The  $NEC_{\text{density}}$  reflects normalized effective counts distributed within the subject body and represents count statistics per subject volume including lung area. The  $NEC_i$  is calculated as shown in Eq. 6, and  $V$  [ $\text{cm}^3$ ] represents the subject volume within the axial extent to be evaluated ( $i = 1$  to  $I$ ), i.e., from the neck to the abdomen in this guideline.

**Acknowledgments** This work was supported in part by a Grant-in-Aid for Cancer Research (25-A-13) from the Japanese Ministry of Health, Labour and Welfare. No conflict of interest exists.

## References

1. Fukukita H, Senda M, Terauchi T, Suzuki K, Daisaki H, Matsumoto K, et al. Japanese guideline for the oncology FDG-PET/CT data acquisition protocol: synopsis of Version 1.0. *Ann Nucl Med.* 2010;24:325–34.
2. Mizuta T, Senda M, Okamura T, Kitamura K, Inaoka Y, Takahashi M, et al. NEC density and liver ROI S/N ratio for image quality control of whole-body FDG-PET scans: comparison with visual assessment. *Mol Imaging Biol.* 2009;11:480–6.
3. National Electrical Manufacturers Association. NEMA standards publication NU 2-2007: performance measurement of positron emission tomographs. Rosslyn: National Electrical Manufacturers Association; 2007.
4. Mawlawi O, Podoloff DA, Kohlmyer S, Williams JJ, Stearns CW, Culp RF, et al. Performance characteristics of a newly developed PET/CT scanner using NEMA standards in 2D and 3D modes. *J Nucl Med.* 2004;45:1734–42.
5. Surti S, Karp JS. Imaging characteristics of a 3-dimensional GSO whole-body PET camera. *J Nucl Med.* 2004;45:1040–9.
6. Erdi YE, Nehmeh SA, Mulnix T, Humm JL, Watson CC. PET performance measurements for an LSO-based combined PET/CT scanner using the National Electrical Manufacturers Association NU 2-2001 standard. *J Nucl Med.* 2004;45:813–21.
7. Teräs M, Tolvanen T, Johansson JJ, Williams JJ, Knuuti J. Performance of the new generation of whole-body PET/CT scanners: discovery STE and discovery VCT. *Eur J Nucl Med Mol Imaging.* 2007;34:1683–92.
8. Kemp BJ, Kim C, Williams JJ, Ganin A, Lowe VJ. NEMA NU 2-2001 performance measurements of an LYSO-based PET/CT system in 2D and 3D acquisition modes. *J Nucl Med.* 2006;47:1960–7.
9. Mejia AA, Nakamura T, Masatoshi I, Hatazawa J, Masaki M, Watanuki S. Estimation of absorbed doses in humans due to intravenous administration of fluorine-18-fluorodeoxyglucose in PET studies. *J Nucl Med.* 1991;32:699–706.
10. Hentschel M, Brink I. Lean body mass-based standardized uptake value, derived from a predictive equation, might be misleading in PET studies. *Eur J Nucl Med Mol Imaging.* 2002;29:1630–8.
11. National Electrical Manufacturers Association. NEMA standards publication NU 2-2001: performance measurement of positron emission tomographs. Rosslyn: National Electrical Manufacturers Association; 2001.