Japanese Journal of Radiology (2018) 36:686–690
https://doi.org/10.1007/s11604-018-0780-x

SPECIAL REPORT

Standardization of [F-18]FDG PET/CT for response evaluation by the Radiologic Society of North America-Quantitative Imaging Biomarker Alliance (RSNA-QIBA) profile: preliminary results from the Japan-QIBA (J-QIBA) activities for Asian international multicenter phase II trial

Hyeyeol Bae[1] · Junichi Tsuchiya[1] · Takehito Okamoto[2] · Ikuko Ito[2] · Yusuke Sonehara[3] · Fumiko Nagahama[4] · Kazunori Kubota[1] · Ukihide Tateishi[1,5]

Received: 27 July 2018 / Accepted: 19 September 2018 / Published online: 24 September 2018

Abstract

Purpose In an Asian international multicenter phase II trial conducted in patients with peripheral T-cell lymphoma (PTCL), [F-18]FDG-PET/CT was used for evaluation of the therapeutic response. Standardization of the PET/CT scanners was necessary before patient enrollment. We therefore standardized the scanners by phantom tests based on the profile approved by the Quantitative Imaging Biomarkers Alliance (QIBA) of Radiological Society of North America (RSNA).

Materials and methods The tests were conducted on 12 scanners in 12 facilities in compliance with the QIBA Profile and used National Electrical Manufacturers Association (NEMA) International Electrotechnical Commission (IEC) body phantoms. We measured three parameters (standardized uptake value [SUV], resolution and noise) and adjusted the imaging parameter values. The indexes recommended in the Japanese Society of Nuclear Medicine (JSNM) guideline were also evaluated.

Results In a total of 12 facilities, 6 facilities required no change in imaging conditions and 6 facilities required changes in imaging parameters. After revision, the three measurements (SUV, resolution and noise) met QIBA criteria at all sites, but 10 of the 12 scanners did not meet JSNM criteria.

Conclusion We standardized imaging conditions using phantoms as required in the RSNA-QIBA profile for response evaluation by [F-18]FDG PET/CT images in a multicenter study.

Keywords Standardization · PET/CT · RSNA · QIBA · J-QIBA

Introduction

Malignant lymphoma consists of various histologic subtypes and can be divided into Hodgkin's and non-Hodgkin's lymphoma. Diagnosis is based on the WHO classification, and selection of treatment and prognosis depend on histologic subtypes [1]. Peripheral T-cell lymphoma (PTCL) is one of the non-Hodgkin's lymphomas and originates from mature T-cells. The course of PTCL is clinically aggressive and poorly responsive to therapy [2–4], and thus new more-effective treatment is desired.[F-18]FDG PET/CT is recommended for the evaluation of therapeutic effects on malignant lymphoma subtypes with high [F-18]FDG avidity [5].

Because of bias and variance in the results obtained from clinical images, quantitative and reproducible measures are

✉ Ukihide Tateishi
ttisdrnm@tmd.ac.jp

1 Department of Diagnostic Radiology and Nuclear Medicine, Tokyo Medical and Dental University, Tokyo, Japan

2 Imaging Service Department, MICRON Inc, Tokyo, Japan

3 Clinical Development, Solasia Pharma K.K, Tokyo, Japan

4 Asia Clinical Development, Solasia Pharma K.K, Tokyo, Japan

5 Japan-Quantitative Imaging Biomarker Alliance (J-QIBA), Japan Radiologic Society, Tokyo, Japan

needed to validate specific metrics used in clinical trials. Imaging biomarkers would be validated and reliably measured and can act as meaningful surrogates for evaluation of therapeutic responses in individuals or groups. Analysis of data collected during the qualification step, substantiating performance as a response measure, could be developed into a reliable method [6–11].

The Quantitative Imaging Biomarkers Alliance (QIBA) was set up by the Radiological Society of North America (RSNA) in order to establish quantitative imaging biomarkers by reducing variability in imaging conditions and the imaging environment [12]. QIBA has 18 biomarker committees, and the [F-18]FDG-PET/CT Biomarker Committee has created a profile for response evaluation by [F-18]FDG-PET/CT in the setting of a clinical trial [13]. The profile addresses the need for phantom test standardization to ensure uniform quantitative performance across all scanners and all sites. Standardization has also been attempted in Japan. The Japan Radiological Society (JRS) has established the Japan-QIBA (J-QIBA) in cooperation with the RSNA-QIBA, and the Japanese Society of Nuclear Medicine (JSNM) has created guidelines governing standardization [14].

Darinaparsin (*S*-dimethylarsino-glutathione) is an organic arsenical used for treatment of malignant tumors [15,16]. Its efficacy for PTCL has been studied. The possible effect on PTCL was suggested by the results of a multicenter phase II study of darinaparsin in patients with relapsed or refractory Hodgkin and non-Hodgkin lymphoma [17]. Darinaparsin was shown to act via the MAPK pathway in a study using lymphoma cells and xenografts in SCID mice [18]. In an Asian international multicenter phase II trial, [F-18] FDG PET/CT scanners in the facilities of all countries were standardized in advance using the [F-18]FDG PET/CT profile developed by the RSNA-QIBA. We show the results of standardization using phantom tests described in the RSNA-QIBA profile.

Materials and methods

Darinaparsin (*S*-dimethylarsino-glutathione) has been evaluated for PTCL in previous studies [17,18]. In a phase II clinical trial of darinaparsin monotherapy in patients with relapsed or refractory PTCL in Asian countries, including South Korea, Taiwan, Hong Kong, and Japan, it was decided that the therapeutic response would be evaluated by central assessment of PET/CT imaging data. Since various PET/CT scanners in each facility would be used, standardization of the scanners was required. As J-QIBA activities, phantom tests of individual PET/CT scanners were performed at all facilities participating in the Darinaparsin Phase II clinical trial before patient enrollment. All the facilities are listed in Table 1. The institutional review board at each center approved the clinical trial.

Table 1 The facilities where phantom tests were conducted

Facility	Location
Samsung Medical Center	South Korea
National Cancer Center	South Korea
Asan Medical Center	South Korea
Severance Hospital	South Korea
Korea Cancer Center Hospital	South Korea
Chonnam National University Hwasun Hospital	South Korea
National Taiwan University Hospital	Taiwan
Taipei Veterans General Hospital	Taiwan
Chang Gung Memorial Hospital Linkou	Taiwan
China Medical University Hospital	Taiwan
National Cheng Kung University Hospital	Taiwan
Hong Kong Integrated Oncology Centre	Hong Kong

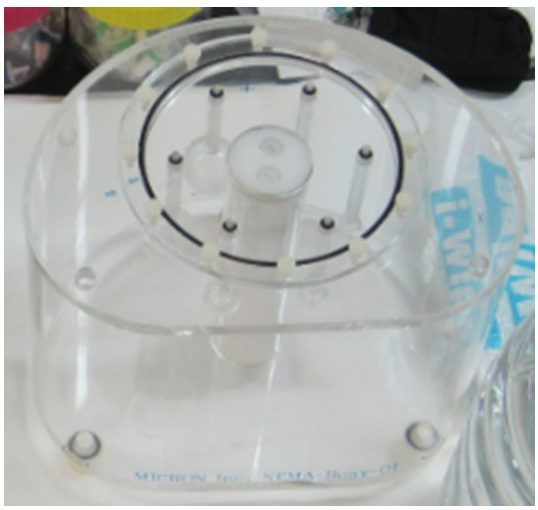

Fig. 1 This was the phantom used for standardization, which had six hot spheres (10, 13, 17, 22, 28 and 37 mm)

The phantom tests were conducted in compliance with the QIBA profile requirements and as stated in *[F-18]FDG-PET/CT as an Imaging Biomarker Measuring Response to Cancer Therapy* [13]. The National Electrical Manufacturers Association (NEMA) International Electrotechnical Commission (IEC) body phantom and [F-18]FDG prepared at each site were used for the phantom test (Fig. 1). The radio activity level was maintained at 3.7–7.2 kBq/ml in the background area of the phantom, and at four times the background level in the hot sphere. Continuous PET data were acquired over a 1–10 min period, and each image was reconstructed with an adequate method and parameters, which were adjusted from default values as necessary at each facility.

To evaluate the scanner, we measured the following three parameters: (a) standardized uptake value (SUV), (b) resolution, and (c) noise as described by QIBA. These measurements were used to assess whether (a) the SUV for the region of interest (ROI) set in the phantom was 1.0 ± 0.1 (b) the 13-mm hot sphere in the phantom was visible, and (c) the coefficient of variation (COV) of the voxel values within the region in the background area was below 15%. Axial uniformity, also mentioned in the profile, was not measured because the shape of the NEMA body phantom was not suitable for the measurement. We evaluated whether these criteria were fulfilled by parameter adjustment.

To clarify the difference between the RSNA-QIBA profile and the guideline in Japan, we also evaluated indexes recommended in *Japanese guideline for the oncology FDG-PET/CT data acquisition protocol* (JSNM guideline) [14], i.e., phantom noise equivalent count ($NEC_{phantom}$), % background variability ($N_{10\,mm}$), % contrast ($Q_{H,\,10\,mm}$), and relative recovery coefficient (RC).

Results

Twelve facilities in Asia (South Korea, Taiwan, and Hong Kong) were enrolled in this trial (Table 1), and standardization was carried out for 12 scanners including the Discovery PET/CT 600, Discovery PET/CT 690, Discovery PET/CT 710, Discovery STE 16, and Discovery VCT (GE Healthcare, total number of scanners: 8) and the TruePoint Biograph 6, TruePoint Biograph 40, Biograph mCT, and Biograph mCT Flow 40-4R (Siemens Healthineers, total number of scanners: 4). The scanners, injected doses, and imaging parameter values for each site are shown in Table 2. At each center, [F-18]FDG was injected in daily practice at 3.7–7.4 MBq/kg or 370 MBq. Scan duration remained in the range of 1.5–3.5 min at all sites except one that used flow motion. We adjusted imaging conditions to meet the criteria approved by the RSNA-QIBA as needed. Change in one or more parameters was needed at 6 of the 12 facilities but not at the other 6 facilities.

In accord with the QIBA profile, the data from the phantom tests for SUV, resolution, and noise were analyzed.

Table 2 List of scanners, injected doses, scan durations, image reconstruction parameters

Site	Scanner	Injected dose	Scan duration (min)		Image reconstruction parameters	
			Initial	Revised	Initial parameters	Revised parameters
A	Discovery PET/CT 710	5.18 MBq/kg	2.0	2.0	3D-OSEM + PSF (Iter: 4, Sub: 18, GF: 4 mm)	3D-OSEM (Iter: 3, Sub: 18, GF: 4 mm)
B	Discovery PET/CT 690	5.92 MBq/kg	2.0	2.0	3D-OSEM + TOF (Iter: 2, Sub: 16, GF: 6.4 mm)	No change
C	TruePoint Biograph 6	5.3 MBq/kg	3.5	3.5	FORE + OSEM (Iter: 2, Sub: 8, GF: 4 mm)	3D-OSEM (Iter: 3, Sub: 21, GF: 6 mm)
D	Discovery STE 16	5.0 MBq/kg	2.5	2.5	3D-OSEM (Iter: 2, Sub: 20, GF: 4.29 mm)	No change
E	TruePoint Biograph 40	3.7 MBq/kg	2.5	2.5	3D-OSEM (Iter: 3, Sub: 8, GF: 4 mm)	3D-OSEM (Iter: 3, Sub: 21, GF: 6 mm)
F	Biograph mCT	370 MBq	1.5	1.5	3D-OSEM + PSF+ TOF (Iter: 2, Sub: 21, GF: 3 mm)	3D-OSEM + TOF (Iter: 2, Sub: 21, GF: 3 mm)
G	Discovery PET/CT 600	3.7 MBq/kg	2.0	2.0	3D-OSEM (Iter: 2, Sub: 16, GF: 6.4 mm)	No change
H	Discovery PET/CT 710	370 MBq	3.0	3.0	QCFX-S ($\beta = 300$)	3D-OSEM + TOF (Iter: 2, Sub: 16, GF: 6 mm)
I	Biograph mCT Flow 40–4R	370 MBq	Motion flow	1.5	3D-OSEM + TOF + PSF (Iter: 2, Sub: 21, GF: 5 mm)	3D-OSEM + TOF (Iter: 2, Sub: 21, GF: 5 mm)
J	Discovery VCT	5.29 MBq/kg	2.5	2.5	3D-OSEM (Iter: 2, Sub: 28, GF: 6 mm)	No change
K	Discovery PET/CT 710	5.18 MBq/kg	2.5	2.5	3D-OSEM + TOF (Iter: 2, Sub: 24, GF: 6.4 mm)	No change
L	Discovery PET/CT 710	7.4 MBq/kg	2.0	2.0	3D-OSEM + TOF (Iter: 3, Sub: 18, GF: 4 mm)	No change

Injected dose is the [F-18]FDG dose injected in daily practice.

FORE Fourier rebinning, *GF* Gaussian filter, *Iter* iteration, *OSEM* ordered subsets expectation maximization, *PSF* point spread function, *Sub* subset, *TOF* time-of-flight

Table 3 List of SUVs, resolutions, and coefficients of variation (COV)

Site	SUV measurement		Resolution measurement	Noise measurement, COV (%)
	Mean	SD		
A	1.0	0.1	Yes	11.9
B	1.0	0.1	Yes	5.2
C	0.9	0.1	Yes	6.1
D	1.0	0.1	Yes	7.7
E	1.1	0.1	Yes	7.1
F	1.1	0.1	Yes	11.7
G	0.9	0.1	Yes	8.3
H	1.1	0.1	Yes	8.6
I	1.0	0.1	Yes	7.9
J	1.1	0.1	Yes	8.7
K	1.0	0.1	Yes	9.7
L	1.0	0.1	Yes	10.4

"SUV measurement" refers to the mean SUV for the ROI set in the phantom, which should be 1.0 ± 0.1

"Resolution measurement" refers to the visibility of the 13-mm hot spheres in the phantom (yes indicates visible)

"Noise measurement" refers to the COV of the voxel values within the region in the background area and should be below 15%

Table 4 Mean and range for the indexes in the JSNM guideline

Index	Mean (Range)
$NEC_{phantom}$ (Mcounts)	13.5 ± 3.3 (8.9–21.6)
$N_{10\ mm}$ (%)	6.2 ± 0.9 (4.1–7.2)
$Q_{H,\ 10\ mm}$ (%)	23.5 ± 7.5 (11.5–38.1)
$Q_{H,\ 10\ mm}/N_{10\ mm}$	3.8 ± 1.1 (2.3–6.0)
$RC_{10\ mm}$	0.51 ± 0.09 (0.39–0.66)

$NEC_{phantom}$ phantom noise equivalent count, $N_{10\ mm}$ % background variability, $Q_{H,\ 10\ mm}$ % contrast, $RC_{10\ mm}$ relative recovery coefficient for 10-mm spheres

SUV for the ROI set in the phantom ranged from 0.9 to 1.1, and the 13-mm spheres in the phantom were visible on all scanners. The maximum COV of the voxel values was 11.9%, which should be below 15% according to the profile. We confirmed that the image quality met all three criteria at all sites after parameter adjustment (Table 3).

After the revision of imaging parameters, we assessed the physical indexes mentioned in the Japanese guideline (JSNM guideline) (Table 4). Ten of the 12 scanners did not meet JSNM criteria ($NEC_{phantom} > 10.8$ Mcounts, $N_{10\ mm} < 5.6\%$, $Q_{H,10mm}/N_{10\ mm} > 2.8$, $RC_{10\ mm} > 0.38$). Patient enrollment began after individual institutions received the results of this field data analysis.

Discussion

This study demonstrated the first attempt as J-QIBA activities to standardize the image quality of [F-18]FDG PET/CT scans used for the evaluation of therapeutic effect in an Asian international multicenter phase II trial using the [F-18]FDG PET/CT profile approved by the RSNA-QIBA [13]. The results of standardization using phantom tests recommended by the RSNA-QIBA showed that image quality standardization was achieved safely and reliably before proceeding to patient enrollment.

In this study, phantom tests were performed at 12 institutions in Asia (South Korea, Taiwan, and Hong Kong). Because of the sufficiently high [F-18]FDG concentration used at each site, scan time extension from that used in the initial protocol was not needed at all sites except one that used flow motion.

Point spread function (PSF) correction was used in the reconstruction process in some facilities as a default but was not adopted in the revision process. In this clinical trial, [F-18]FDG PET/CT was used for evaluation of the therapeutic effect rather than detection, and it was important to minimize the differences between the scanners. PSF correction is known to increase noise and Gibbs artifact [19,20] and not considered appropriate.

In addition to the measurements stated in the QIBA profile, the parameters for image quality required by *Japanese guideline for the oncology FDG-PET/CT data acquisition protocol* (JSNM guideline) were also assessed [14]. Standardization by phantom tests as recommended in the JSNM guideline for a clinical trial was reported previously [21] and evaluation was done in this study in the same way. The imaging conditions were adjusted to meet the criteria in the QIBA profile at all the facilities, but 10 of the 12 scanners could not fulfill the recommendations when evaluated by the JSNM guideline. This result indicated criteria for image quality referred to in the QIBA profile were easier to be met than those mentioned in the JSNM guideline, and the standardization procedure in the QIBA profile was regarded as convenient because of the ease of phantom test introduction into the international multicenter study.

Standardization procedure in this study was based on (but not fully compliant with) the one described in the QIBA profile. Axial uniformity measurement, one of the criteria used in phantom testing according to the QIBA profile, could not be assessed because of the shape of the NEMA body phantom. Furthermore, applying harmonization strategies might be required for enigmatic studies in the future.

Quantitative and reproducible measures from imaging studies are needed to validate specific metrics in clinical trials and clinical practice because bias and variance in

the results obtained from clinical images can come from several sources. RSNA-QIBA has developed a flexible framework to organize the work of its coordinating and biomarker committees, which is to identify reproducible quantitative imaging biomarkers. The RSNA-QIBA has made an effort to liaison with the European Imaging Biomarkers Alliance (EIBALL) and J-QIBA. As shown by the initial results of an Asian international multicenter phase II trial as J-QIBA activities, results obtained in Asia are possible to correspond to those of international clinical trials in western countries using the RSNA-QIBA profile.

In conclusion, using the RSNA-QIBA profile, we standardized imaging conditions by phantom tests for response evaluation by [F-18]FDG PET/CT images acquired in a multicenter study. J-QIBA can settle quantitative imaging data of Asian international studies.

Acknowledgements We gratefully appreciate the cooperation of the following members and institutions, JoonYoung Choi (Samsung Medical Center), SeokKi Kim (National Cancer Center), Jin Suck Ryu, JeongSu Oh (Asan Medical Center), MiJin Yun (Severance Hospital), Byeong-Il Kim (Korea Cancer Center Hospital), JungJun Min, SeongYoung Kwon (Chonnam National University Hwasun Hospital), Ruoh-Fang Yen (National Taiwan University Hospital), Wen-Sheng Huang, Bang-Hung Yang (Taipei Veterans General Hospital), Chien Lin, Wen-Cheng Huang (Chang Gung Memorial Hospital Linkou), Te-Chun Hsieh (China Medical University Hospital), Nan-Tsing Chiu (National Cheng Kung University Hospital), and Yang Kwok Wai, Michael (Hong Kong Integrated Oncology Centre). We have also been given helpful suggestions by Edward Jackson, PhD and Alexander Guimaraes, MD, PhD, Steering Committee, Quantitative Imaging Biomarkers Alliance (QIBA), Radiological Society of North America (RSNA).

Funding No funding was received by any of the authors.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, Thiele J, editors. WHO classification of Tumours of Haematopoietic and Lymphoid Tissues. Revised. 4th ed. Lyon: International Agency for Research on Cancer; 2017.
2. Armitage JO, Vose JM, Weisenburger DD. Towards understanding the peripheral T-cell lymphomas. Ann Oncol. 2004;15:1447–9.
3. Rodriguez-Abreu D, Filho VB, Zucca E. Peripheral T-cell lymphomas, unspecified (or not otherwise specified): a review. Hematol Oncol. 2008;26:8–20.
4. Chihara D, Fanale MA, Miranda RN, Noorani M, Westin JR, Nastoupil LJ, et al. The survival outcome of patients with relapsed/refractory peripheral T-cell lymphoma-not otherwise specified and angioimmunoblastic T-cell lymphoma. Br J Haematol. 2017;176:750–8.
5. Cheson BD, Fisher RI, Barrington SF, Cavalli F, Schwartz LH, Zucca E, et al. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. J Clin Oncol. 2014;32:3059–68.
6. Sargent DJ, Rubinstein L, Schwartz L, Dancey JE, Gatsonis C, Dodd LE, et al. Validation of novel imaging methodologies for use as cancer clinical trial end-points. Eur J Cancer. 2009;45:290–9.
7. Meyer CR, Armato SG, Fenimore CP, McLennan G, Bidaut LM, Barboriak DP, et al. Quantitative imaging to assess tumor response to therapy: common themes of measurement, truth data, and error sources. Transl Oncol. 2009;2:198–210.
8. Buckler AJ, Bresolin L, Dunnick NR, Sullivan DC. A collaborative enterprise for multi-stakeholder participation in the advancement of quantitative imaging. Radiology. 2011;258:906–14.
9. Buckler AJ, Bresolin L, Dunnick NR, Sullivan DC, Aerts HJ, Bendriem B, et al. Quantitative imaging test approval and biomarker qualification: interrelated but distinct activities. Radiology. 2011;259:875–84.
10. Sullivan DC, Obuchowski NA, Kessler LG, Raunig DL, Gatsonis C, Huang EP, et al. Metrology standards for quantitative imaging biomarkers. Radiology. 2015;277:813–25.
11. O'Connor JP, Aboagye EO, Adams JE, Aerts HJ, Barrington SF, Beer AJ, et al. Imaging biomarker roadmap for cancer studies. Nat Rev Clin Oncol. 2017;14:169–86.
12. Quantitative Imaging Biomarkers Alliance. Radiological Society of North America. https://www.rsna.org/QIBA/
13. FDG-PET, CT,. Technical Committee. FDG-PET, CT,. as an Imaging Biomarker Measuring Response to Cancer Therapy, Quantitative Imaging Biomarkers Alliance, Version 1.11, Publicly Reviewed Version. QIBA, November 10, 2016. Available from: https://www.rsna.org/QIBA/
14. Fukukita H, Suzuki K, Matsumoto K, Terauchi T, Daisaki H, Ikari Y, et al. Japanese guideline for the oncology FDG-PET/CT data acquisition protocol: synopsis of Version 2.0. Ann Nucl Med. 2014;28:693–705.
15. Hirano S, Kobayashi Y. Cytotoxic effects of S-(dimethylarsino)-glutathione: a putative intermediate metabolite of inorganic arsenicals. Toxicology. 2006;227:45–52.
16. Matulis SM, Morales AA, Yehiayan L, Croutch C, Gutman D, Cai Y, et al. Darinaparsin induces a unique cellular response and is active in an arsenic trioxide-resistant myeloma cell line. Mol Cancer Ther. 2009;8:1197–206.
17. Hosein PJ, Craig MD, Tallman MS, Boccia RV, Hamilton BL, Lewis JJ, et al. A multicenter phase II study of darinaparsin in relapsed or refractory Hodgkin's and non-Hodgkin's lymphoma. Am J Hematol. 2012;87:111–4.
18. Ravi D, Bhalla S, Gartenhaus RB, Crombie J, Kandela I, Sharma J, et al. The novel organic arsenical darinaparsin induces MAPK-mediated and SHP1-dependent cell death in T-cell lymphoma and Hodgkin lymphoma cells and human xenograft models. Clin Cancer Res. 2014;20:6023–33.
19. Lee YS, Kim JS, Kim KM, Kang JH, Lim SM, Kim HJ. Performance measurement of PSF modeling reconstruction (True X) on Siemens Biograph TruePoint TrueV PET/CT. Ann Nucl Med. 2014;28:340–8.
20. van der Vos CS, Koopman D, Rijnsdorp S, Arends AJ, Boellaard R, van Dalen JA, et al. Quantification, improvement, and harmonization of small lesion detection with state-of-the-art PET. Eur J Nucl Med Mol Imaging. 2017;44(Suppl 1):4–16.
21. Daisaki H, Tateishi U, Terauchi T, Tatsumi M, Suzuki K, Shimada N, et al. Standardization of image quality across multiple centers by optimization of acquisition and reconstruction parameters with interim FDG-PET/CT for evaluating diffuse large B cell lymphoma. Ann Nucl Med. 2013;27:225–32.