The need of a system phantom for quantitative hybrid nuclear imaging of PET/CT: A systematic review

Muhammad Hafiz Hanafi, MSc[1], Noramaliza Mohd Noor, PhD[2,3], Fathinul Fikri Ahmad Saad, MMed[2,3], Saiful Hasley Ramli, PhD[4], Faizal Mohamed, PhD[5], Marianne Musarudin, PhD[6]

[1]Centre for Diagnostic Nuclear Imaging, Universiti Putra Malaysia, [2]Department of Radiology, Faculty Medicine and Health Sciences, Universiti Putra Malaysia, [3]Department of Radiology, Teaching Hospital Universiti Putra Malaysia, Universiti Putra Malaysia, [4]Department of Industrial Design, Faculty of Design and Architecture, Universiti Putra Malaysia, [5]Department of Applied Physics, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, [6]Medical Radiation Programme, School of Health Sciences, Universiti Sains Malaysia

ABSTRACT

Introduction: There has already been a rising demand in utilising phantom for hybrid Positron Emission Tomography/ Computed Tomography (PET/CT) scanner of nuclear imaging. This review further clarifies this topic and investigates how the previous research phantoms operated with the need for quantitative hybrid nuclear imaging of PET/CT while providing a relatively high image quality when it was performed. In this article, the necessity of previous and current phantom studies in hybrid nuclear imaging of PET/CT scanners is reviewed.

Methods: PubMed and Google Scholar were systematically searched for the relevant studies by following the PRISMA 2009 checklist. A past decade literature search was conducted from 2010 until November 2020 to secure the relevance of the phantom study. Databases were recruited using keywords such as phantom, quantification, standardisation, harmonisation, image quality, standardised uptake value and multicentre study. However, all keywords were related to PET/CT. All abstracts and eligible full-text articles were screened independently, and finally, the quality assessments of this review were performed.

Results: From the 200 retrieved articles, 80 were rejected after the screening of the abstracts and 35 after reading the full-text. The 20 accepted articles addressed the distribution of phantom types used in selected articles studies which were NEMA (67%), ACR (8%) and others (25%). The articles showed the various experimental studies, either phantom studies (35%) or phantom plus clinical studies (65%). For clinical studies (n = 829), the distribution of prospective studies was (n = 674) and retrospective studies was (n =155). The distribution of phantom pathway application showed the studies focused on 40% of reconstruction protocol studies, 30% of the multicentre and standardisation of accreditation program studies, and 30% of the quantification of uptake values studies.

Conclusions: According to this review, the phantom study have a pivotal role in hybrid nuclear imaging of PET/CT either in technical aspects of the scanners (such as data acquisition and reconstruction protocol) or clinical characteristics of patients. In addition to this, the necessity to identify the suitable system phantoms to use within PET/CT scans by considering the continuous development of new phantom studies are needed. Researchers are encouraged to adopt efforts on phantom quantitative validation, including verification with clinical data of patients.

KEYWORDS:
PET/CT phantom, quantification, standardisation, harmonisation, image quality

INTRODUCTION

Over the past decade, there has already been a rising demand in utilising phantom for hybrid Positron Emission Tomography/ Computed Tomography (PET/CT) scanner of nuclear imaging. PET/CT scanners perform a significant role in contemporary nuclear imaging as an outcome of their hybrid existence. A Hybrid PET/CT scanner can show the information of the image by merging metabolic imaging (PET) and morphological imaging with computed tomography (CT).[1]

Phantom is commonly used as a PET/CT scanner validation routine in the quality control (QC) process. The quality control process is obligatory to validate quantitative PET/CT imaging in clinical practice. The phantom can be used for acceptance analysis, routine consistency measurement, precision testing of reconstructed image quality, simulated evaluation of whole-body imaging, identification of non-uniform artefacts, and further evaluation testing.[2]

The changing performance caused by different eras of PET/CT scanners like time-of-flight (TOF), two-dimensional (2D) and three-dimensional (3D) acquisitions technology lead to different quantitatively of image quality. Moreover, the problems also include diverse reconstruction technologies such as point spread function (PSF), or Bayesian penalised-likelihood (BPL) and resolution recovery reconstruction. Multicentre standards of PET/CT systems must not be based on the minimum performing scanners. However, they are required to sustain the maximum standard in the performance of the scanners by implementing further evaluation testing parameters using phantom as one of the tools.[3]

This article was accepted: 08 May 2021
Corresponding Author: Muhammad Hafiz Hanafi
Email: muhammadhafizhanafi90@gmail.com

This review further clarifies this topic and investigates on how the previous research phantoms operated with the need for quantitative hybrid nuclear imaging of PET/CT while providing a relatively high image quality when it was performed. Our review is expected to identify the following questions. This purpose of the review is to respond to the following inquiries: (i) What is the particular phantom used in PET/CT? (ii) What are the potential benefits of requiring a phantom study in PET/CT scan? (iii) Is utilising the phantom just for quality control purposes only?

MATERIALS AND METHODS

Search of the past decade literature review was conducted to secure the relevance of the phantom studies of PET/CT by following the checklist of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses 2009 (PRISMA 2009).[4]

Eligibility criteria

Our inclusion criteria included all phantoms studies related to hybrid nuclear imaging of PET/CT. The exclusion criteria for this study were:

- The phantoms studies were related to imaging of CT, MRI, SPECT, ultrasound, or other imaging modalities.
- The articles were not related to PET/CT phantom studies.
- The studies were related to animal trials.
- The articles were not a decade published (Articles published before 2010).
- The articles have not been published in peer-reviewed journals.

Search strategies and information sources

We conducted a systematic global analysis of various electronic databases (including PubMed and Google Scholar). Since phantoms studies in PET/CT are comparatively new in imaging applications studies in nuclear imaging, the literature search was organised from 2010 until November 30, 2020. Databases were recruited using keywords phrase searching and Medical Subject Headings (MeSH) methods such as "phantom*", "quantification*", "standardisation*", "harmonisation*", "image quality*", "standardised uptake value*", "multicentre study*". However, all keywords were required to be related to PET/CT or PET. The searches were confined to articles in the English language only. Reference lists of related publications were also identified.

Data extraction for study selection

Search strategies were applied with initial findings imported and integrated into the Mendeley Desktop, the reference management platform (version 1.19.4 ©2008-2018 Mendeley Ltd, Elsevier). After eliminating duplicates, the remaining titles and abstracts were reviewed for inclusion. Full texts were extracted and manually assessed from the relevant articles.

Using structured data extraction techniques, the authors retrieved and compared the data separately. Discrepancies were explored. Data obtained included the published year, type of phantom used, clinical approaches in phantom research, clinical design experiments, phantom pathway study, and outcomes of the studies.

By thoroughly reviewing the aims of the studies, all studies obtained from the search procedure were checked if they were within the scope of the current research or not. The authors of this paper independently reviewed the described and chosen articles. They also reviewed and prepared detailed notes to outline the research purposes, methodology, techniques, significant findings and recommendations and made a definitive decision based on research criteria.

Fifty-five reports were found of articles in full-text. Thirty-five articles were not considered since the papers were not related to phantoms experimental and were excluded for reasons. As a result, the search strategy identified twenty articles according to our study criteria. The review process for related publications using the PRISMA recommendation is shown in Figure 1.

Risk of bias

No qualitative score was applied for the selection of the study. Data obtained from reviewed articles were structured to minimise potential biases using other reviewers only in queries. Data extracted from research that did not follow their possible standardisation requirements were excluded and withdrawn.

Summary data analysis

Descriptive statistics was used to sum up, the information with percentage values for dichotomous variables. Percentages were calculated to determine the per cent of the sample corresponding to the specified frequency. The values are usually presented without decimal points and significant figures (according to American Psychological Association© year 2020 seventh edition standards).

RESULTS

Literature characteristics

There were 20 full articles identified from the year 2010 until 2020 that focused on the experimental phantom studies (Table I). The articles that reviewed the literature were detailed into four information groups (subjects, type of phantom, the aim of the study, and quality of research finding) as demonstrated in Table I.

Distribution of phantom types

This review identified various phantom types used for the PET/CT studies, as presented in Figure 2. First, the phantom used was the NEMA phantom, which was designed according to the endorsement of the National Electrical Manufacturers Association (NEMA). The second was the Jaszczak Deluxe Flangeless PET phantom, which was designed according to the requirement of the American College of Radiology (ACR). Simultaneously, the rest was classified as other phantoms (anthropomorphic, cylindrical and modified micro hollow).

There were 16 articles reported using NEMA phantom as their study in PET/CT. As for ACR phantom, only two phantoms and six papers used other phantoms as their study. The majority used NEMA phantom with 67% in percentage distribution, ACR phantom 8%, and other phantoms were 25%.

Distribution of phantom studies
Table I shows the distribution of PET/CT studies either in the only phantom study or phantom-clinical study approaches. The only phantom study is the research that has been conducted using phantom only and without clinical approaches. For the phantom-clinical study, the research were worked out using phantom and clinical study either retrospectively or prospectively. The only phantom study percentage was 35%, and the phantom-clinical study was 65% from all the reviewed articles. The phantom-clinical study methods were divided into retrospective and prospective studies, with percentages of 20% and 45%, respectively. In the prospective study, the clinical data were directly collected during the period of research study. However, in the retrospective study, the clinical data were usually sampled and collected from the past or previous PET/CT examination.

Distribution of clinical studies in PET/CT
Table I illustrates a more specific distribution of clinical studies, either retrospective or prospective. For retrospective clinical studies, the total number of clinical data was 155. The data was presented by four reviewed articles, they were Devriese J et al. (n=64), Armstrong I et al. (n=68), Wielaard J et al. (n=15), and te Riet et al. (n=8).

In the prospective studies, the total number of clinical data was 674. Nine reviewed articles studied about prospective clinical studies, which were Makris N et al. (n=10), Kelly M & Declerck J (n=10), Quak E et al. (n=517), Texte E et al. (n=20), Hoetjes N et al. (n=25), Lasnon C et al. (n= 52), Kaalep et al. (n=30), Caribe et al. (n=1) and Kero T et al. (n=9).

Result of phantom pathway study in reviewed articles of PET/CT researches
Figure 3 shows the result of phantom pathway studies in PET/CT researches. The pathway was divided into three groups of study approaches, which were a) quantification of uptake values, b) reconstruction protocols, and c) multicentre studies and accreditation programme.

The reconstruction protocols percentage was 40% and the highest of phantom pathway studies with eight reviewed papers. However, the percentage of quantification of uptake values was 30%. The same goes for multicentre studies & accreditation programme, with 30% with six reviewed papers.

DISCUSSION

This review focuses on the need of a system phantom for quantitative hybrid nuclear imaging of PET/CT. The reviewed articles were restricted to a decade publication to highlight contemporary PET/CT studies' outcomes using various phantoms, either study conducted on phantom only or add-on with clinical research. The current review was undertaken and reported using the recommended PRISMA guidelines.

Application of different types of phantoms in PET/CT studies
The reviewed papers highlighted the quantitative studies of PET/CT's hybrid nuclear imaging of PET/CT using various system phantoms. As reported in Figure 1, the reviewed articles reported that the phantoms used were NEMA, ACR and others (anthropomorphic, cylindrical and modified micro hollow). All the phantoms were specific for PET imaging systems and reliable for nuclear imaging studies, as presented in Table II.

From the reviewed papers' extracted information, NEMA phantoms represented the highest number of phantom studies with a total of 16 reviewed articles. Through all that articles, NEMA phantom studies clarified the research pathway of a) multicentre studies and accreditation programme, b) reconstruction protocols and, c) quantification of uptake values.

Data Spectrum explained that the NEMA phantom is fabricated to standardise the evaluation of PET scanner performance according to the standard of the National Electrical Manufacturers Association. The phantom has a part of the body and the lung part attached with six fillable spheres of different sizes. The phantom provides imaging information particularly through PET as the camera-based coincidence imaging techniques. Besides, the phantom helps with system image quality as well as the accuracy of any corrections used.[24]

Many researchers used NEMA phantom due to the specific useful for clinical study, especially for brain and cardiac imaging studies. Biodex Medical System, Inc. informed that this phantom could determine cardiac and brain imaging's synchronise count rate features. Moreover, it complies with NEMA 2012 standard.[25]

However, ACR and other phantoms only showed two common research applications, either in multicentre studies and accreditation programme, or reconstruction protocols. Only two reviewed papers focused on ACR phantom studies and six reviewed articles reported about other phantoms studies.

Biodex Medical System, Inc. and Supertech, Inc. stated that the ACR phantom offers reliable and accurate performance information for any PET systems. From a single scan of the phantom, it can assess various evaluation characteristics of PET systems. The function of the transverse line spread on-axis and off-axis can be easily determined without removing the cover plate. The ACR phantom for PET meets the requirements according to the standard of the ACR.[26,27]

According to the Report of the American Association of Physicists in Medicine Task Group 126: PET/CT Acceptance Testing and Quality Assurance 2019, ACR phantom can evaluate PET image contrast and scatter attenuation correction. The ACR phantom is characterised by four hot vials of varying diameters with a fixed activity concentration relative to the background and three vials of varying material densities.[28]

The SUV ratios will be used in the calculation. The maximum SUV measurements from the four hot vials will be used to measure image contrast. The scatter/attenuation from the Teflon, air, water, and background regions will be calculated using mean and minimum SUV measurements. These values

Table I: Information of each reviewed article regarding the subjects, type of phantom, aim of study and quality of finding

Author	Subjects	Phantom studies	Aim	Quality of finding
Kaalep A et al. [3]	Phantom & patient (n=30)[P]	NEMA	To study the role of EARL-2 revised accreditation guideline on quantitative measurements of clinical PET/CT studies	The updated EARL-2 recommendations resulted in higher SUVs, lower MATV, and similar TLGs.
Makris NE et al. [5]	Phantom & patient (n=10)[P]	NEMA ACR Anthropomorphic	Determine whether the phantoms are ideally suited to detect variations in image quality and quantification, and the methods to identify volumes of interest (VOI) are the least sensitive to these differences.	The three phantoms investigated in this study were suitable for harmonising various scanner quantitative performances, suggesting more potential for harmonising image quality and quantification.
Kelly MD and Declerck JM [6]	Phantom & patient (n=10)[P]	NEMA	They proposed a new approach, reference Standardised Uptake Value (SUVref), to reduce the quantitative variation arising from reconstruction protocol inconsistencies.	This reduction in variance significantly improves clinical image quantitative comparison to assess the disease's treatment response or progression.
Quak E et al. [7]	Phantom & patient (n=517)[P]	NEMA	To validate a specific software tool (EQ.PET) to harmonise SUVs across various PET systems regardless of the reconstruction algorithm used.	This is mostly applicable to multicentre trials and can provide precise quantification for restaging.
Devriese J et al.[8]	Phantom & patient (n=64)[R]	NEMA	To compare lesion SUV values collected via two different reconstruction protocols: a) GE's latest clinical lesion detection protocol (Q.Clear); b) The EARL harmonisation protocol, using the PERCIST protocol	It is advisable to select the EARL protocol for multicentre studies and individual therapy response evaluation to accurately compare the SUL between patients, scanners, and centres.
Rogasch JMM et al. [9]	Phantom	Cylindrical phantom	To investigate the impact of reconstruction integration on various SBRs.	The use of reconstruction for quantitative PET data should be conducted with caution (if SUV of lesions with high contrast compared to low contrast)
Texte E et al. [10]	Phantom & patient (n=20)[P]	ACR	To assess the BPL reconstruction algorithm's effect compared to OSEM on the hypoxia PET/CT images.	The BPL algorithm explicitly raises the quantitative parameters and contrast on PET/CT reconstruction that is consistent with all other papers studying this reconstruction algorithm.
Akerele MI et al. [11]	Phantom	NEMA	To implement the newly proposed background correction and assess its efficiency in lesion quantification accuracy and comparison, using both simulated and actual clinical PET.	Improved quantification and more precise lesions detection were achieved with the newly proposed background correction technique.
Armstrong IS et al. [12]	Phantom & patient (n=68)[R]	NEMA	To study the effect of PSF and TOF simulation on SUVmax	Gains in SNR were seen in both implementations, with the most significant gains seen for matched SUVmax post-filters.
Wielaard J et al. [13]	Phantom & patient (n=15)[R]	NEMA	Determination of minimum 68Ga-injected operation for clinical PSMA imaging studies	Method indicates that a maximum noise level of 25% is sufficient for the proper analysis and quantification of 68Ga-PSMA studies.
te Riet J et al.[14]	Phantom & patient (n=8)[R]	NEMA modified Micro Hollow Sphere (MHS) phantom	To determine the efficiency of BPL compared to OSEM+PSF in 18F-FDG studies acquired under specific clinical conditions using phantom and patient-based studies.	The efficiency of the BPL algorithm is superior to the standard OSEM+PSF algorithm in small lesion detectability.
Hoetjes NJ et al. [15]	Phantom & patient (n=25)[P]	NEMA	To assess the reliability of several PET-based PVC techniques for oncological whole-body 18F-FDG studies	PVC techniques can be used for more accurate, yet equally precise treatment response assessments.

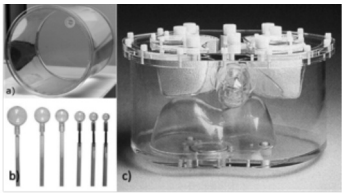
cont..... pg 555

cont from..... pg 554

Author	Subjects	Phantom studies	Aim	Quality of finding
Lasnon C et al. [16]	Phantom & patient (n=52)[P]	NEMA	To analyse a strategy in patients imaged on a PET/CT system equipped with reconstruction of PSF.	This can be used in multicentre trials when using SUV for monitoring therapy or as a diagnostic or prognostic tool. This technique validated in NSCLC patients may be extrapolated to other solid tumours.
Karlberg AM et al. [17]	Phantom	NEMA	To conduct a specific quantitative comparison of the performance of PET image quality between PET/CT and PET/MR	The SUV measurements indicate good agreement for both systems.
Kaalep A et al. [18]	Phantom	NEMA	To report the results and impacts of the accreditation programme on the participating PET/CT systems.	Suggested analysis and upgrade to account for developments in acquisition and reconstruction technologies in PET/CT.
Ho Shon I et al. [19]	Phantom	Anthropomorphic Torso Phantom	To determine the minimum CT acquisition parameters required for maintaining the accuracy of AC for PET reconstruction.	The value of SUV COV was equal at higher CT exposures, regardless of reconstruction algorithm.
Boellaard R et al. 20	Phantom	Cylindrical phantom	To present a standardised imaging method for fixed FDG PET/CT data acquisition, QC and QA.	A new version of the guidelines only addresses combined or integrated whole-body 3D PET/CT systems.
Caribé P et al. [21]	Phantom & patient (n=1)[P]	NEMA	To evaluate different β-factors compared to a clinical post-filter kernel by using BSREM instead of OSEM.	The BSREM reconstruction algorithm provided the opportunity to minimise noise by a factor of 2–4 without a loss of contrast compared to OSEM reconstructions for all data evaluated.
Kaalep A et al. [22]	Phantom	NEMA Cylindrical phantom	To investigate the inconsistency of quantitative performance and the viability of quantitative harmonisation in 89Zr PET/CT imaging.	Harmonisation of PET/CT scanners for quantitative 89Zr studies is feasible when adequate cross-calibration scanner-dose calibration and harmonised image reconstruction procedures are followed. The accreditation programme for PET/CT scanners will support multicentre 89Zr quantitative studies.
Kero T et al. [23]	Phantom & patient (n=9)[P]	NEMA	To validate the simplified methods retention index (RI) and standardised uptake value (SUV). For quantification of cardiac 11C-PIB uptake in amyloidosis.	RI and SUV resulted in a high correlation with quantitative results from this kinetic model, using either individual or population average metabolite data.

(n)P: Prospective clinical study, (n)R: Retrospective clinical study, EARL: The EANM accreditation programme Research 4 Life, SUV: standardised uptake values, MATV: metabolically active tumour volume, TLG: Total lesion glycolysis, VOI: Volume of interest, EQ-PET: Siemens new technique of PET/CT filter, GE: General Electric company, SUL: SUV corrected for lean body mass, PERCIST: Positron Emission Tomography Response Criteria in Solid Tumour, SBR: signal-background ratio, BPL: Bayesian penalised-likelihood, OSEM: the ordered subset expectation maximisation, PSF: point-spread-function, TOF: time-of-flight, SNR: signal-to-noise ratio, PSMA: Prostate-specific membrane antigen, 68Ga: Gallium-68, NEMA: National Electrical Manufacturers Association, ACR: American College of Radiology, PET/CT: Positron Emission Tomography - Computed Tomography, PET/MR: : Positron Emission Tomography - Magnetic Resonance, QA: Quality assurance, QC: Quality control, AC: Attenuation-Corrected, NSCLC: Non-small-cell lung carcinoma, BSREM: Block sequential regularised expectation maximisation, 89Zr: zirconium-89, 11C-PIB: Carbon-11-labeled Pittsburgh Compound-B

Table II: The type of PET/CT phantoms used in the reviewed papers

Type of standard PET/CT phantoms	The description of PET/CT phantoms	The phantoms used in the reviewed papers
	**ACR phantom** Name: Esser Flangeless Deluxe PET Phantom™ This phantom meets the ACR requirements, which provide consistent performance information for PET/CT system. [26,27] This phantom is suitable for system performance evaluations such as collimator, artefacts, calibration, and reconstruction parameters, acceptance testing routine, quality assurance and control. [26,27]	Makris NE et al. [5] Texte E et al. [10]
	**NEMA phantom** Name: PET Phantom - NEMA 2012/IEC 2008 2012 NEMA Standards, ideal for whole-body PET. This phantom complies with NEMA 2012 Standard. Generally used to evaluate reconstructed image quality in whole-body PET and camera-based coincidence imaging and used in research. [24,25]	Kaalep A et al. [3] Makris NE et al. [5] Kelly MD and Declerck JM 6 Quak E et al. [7] Devriese J et al. [8] Akerele MI et al. [11] Armstrong IS et al. [12] Wielaard J et al. [13] te Riet J et al.[14] Hoetjes NJ et al. [15] Lasnon C et al. [16] Karlberg AM et al. [17] Kaalep A et al. [18] Caribé P et al. [21] Kaalep A et al. [22] Kero T et al. [23]
	**Other phantoms** Name: a) uniform fillable cylinder phantom, b) Micro Hollow Sphere (MHS) phantom, c) anthropomorphic phantom These phantoms are suitable for evaluating new image fusion software, evaluating new attenuation correction algorithms and particular specific research.	Makris NE et al. [5] Rogasch JMM et al. [9] te Riet J et al.[14] Ho Shon I et al. [19] Boellaard R et al. [20] Kaalep A et al. [22]

are calculated on reconstructed images with all corrections applied (attenuation, scatter, random counts, dead time and others).[28]

The difference between NEMA phantom and ACR phantom is the difference of fillable vials between them. NEMA phantom has six fillable vials with the shape of a sphere and the volume of the sphere is $4/3\pi r^3$. While ACR only has four fillable vials with the shape of a cylinder and the volume of the cylinder is $\pi r^2 h$. However, NEMA can give more information on uptake value since the image slice of the sphere keeps changing due to the difference in diameter. Compared to the ACR phantom with the vials of the cylinder, its diameter is fixed.

The reviewed articles used the anthropomorphic phantom, cylindrical phantom, and micro hollow sphere phantom for other phantoms. The function of the anthropomorphic phantom is utilised in the assessment of non-uniform attenuation and scatter correction techniques. The phantom has a wide body-shaped cylinder including the mimic parts of liver, lung and spine inserts, that is capable of simulating radioactivity distributions' anatomical structures for the upper torso of average to patients.[29]

Next, a micro hollow sphere phantom specialises in simulating small-scale hot or cold spherical vials and presenting a quantitative evaluation of spatial resolution of small object size effects and reconstruction methods. The phantom can also be utilised to evaluate the uniformity.[30] Finally, yet significantly, the cylindrical phantom is a water-filled cylinder phantom containing a uniform injected radioisotope solution like ^{18}F-FDG, ^{68}Ga or the others. This phantom can be used to assess uniformity.

Application of phantoms in PET/CT quantification studies

Quantification of uptake values has numerous clinical applications, including cancer diagnosis and staging. Especially in ^{18}F-FDG of PET/CT quantification, which portrays a crucial function in diagnosing and staging FDG-avid tumours.[31] Hoetjes NJ et al. demonstrated that the partial volume effects, which results in an increased underestimation of standardised uptake value with decreasing tumour volume, affected the quantitative accuracy and precision.[15] This affects the accuracy and precision of quantification of ^{18}F-FDG PET/CT. Therefore, the simulation and phantom experiments were performed to assess PVC's performance accurately corrected SUV of the

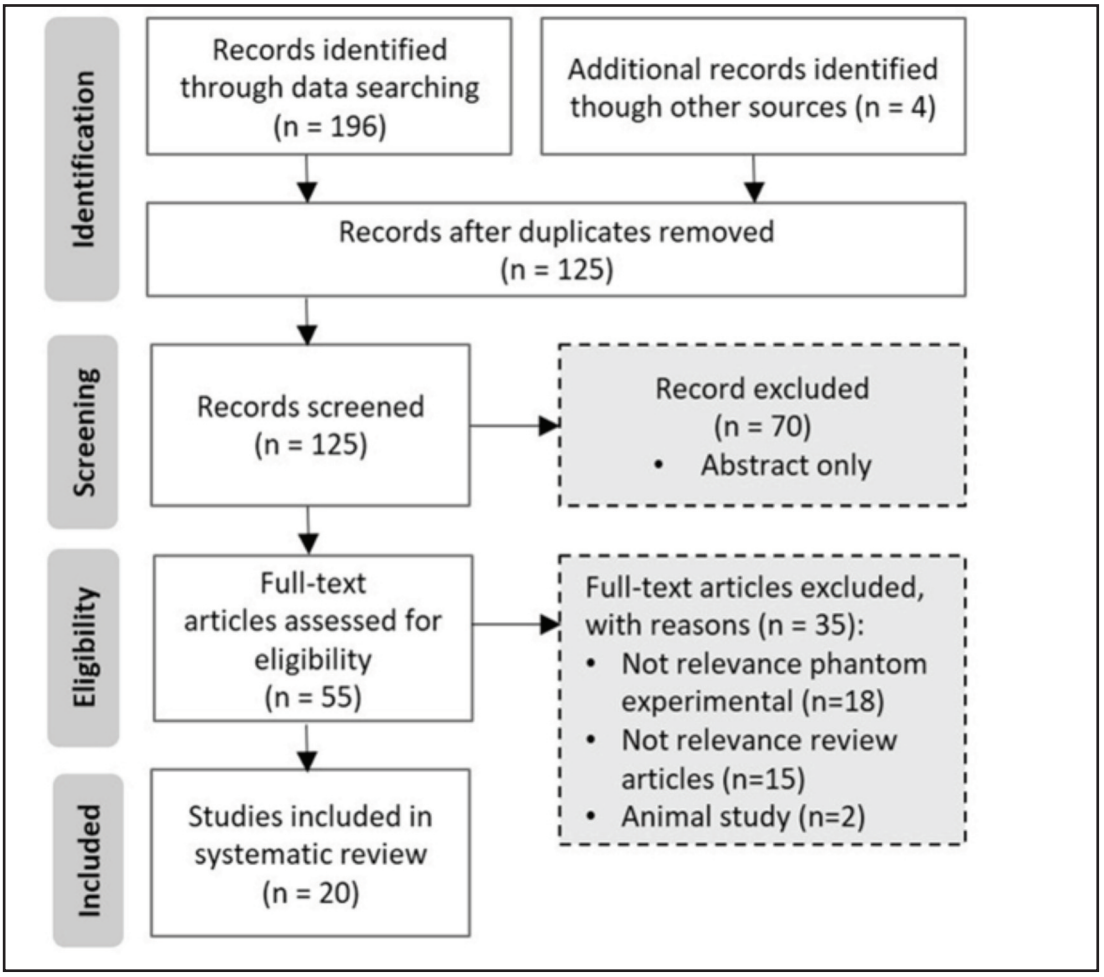


Fig. 1: PRISMA flowchart for PET/CT phantom studies selection process.

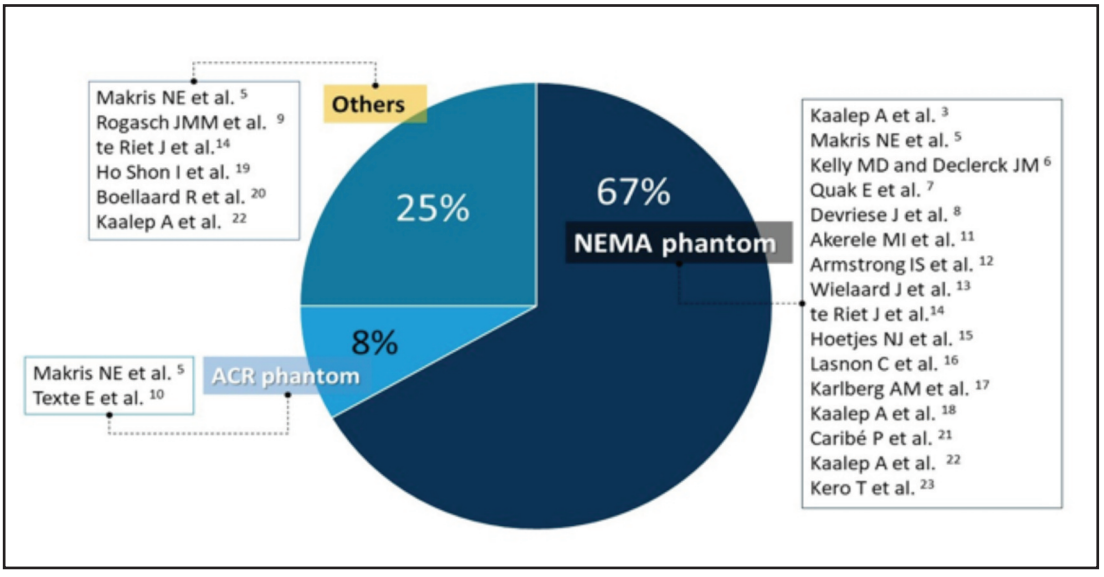


Fig. 2: Distribution of phantoms type used in PET/CT studies in reviewed articles.

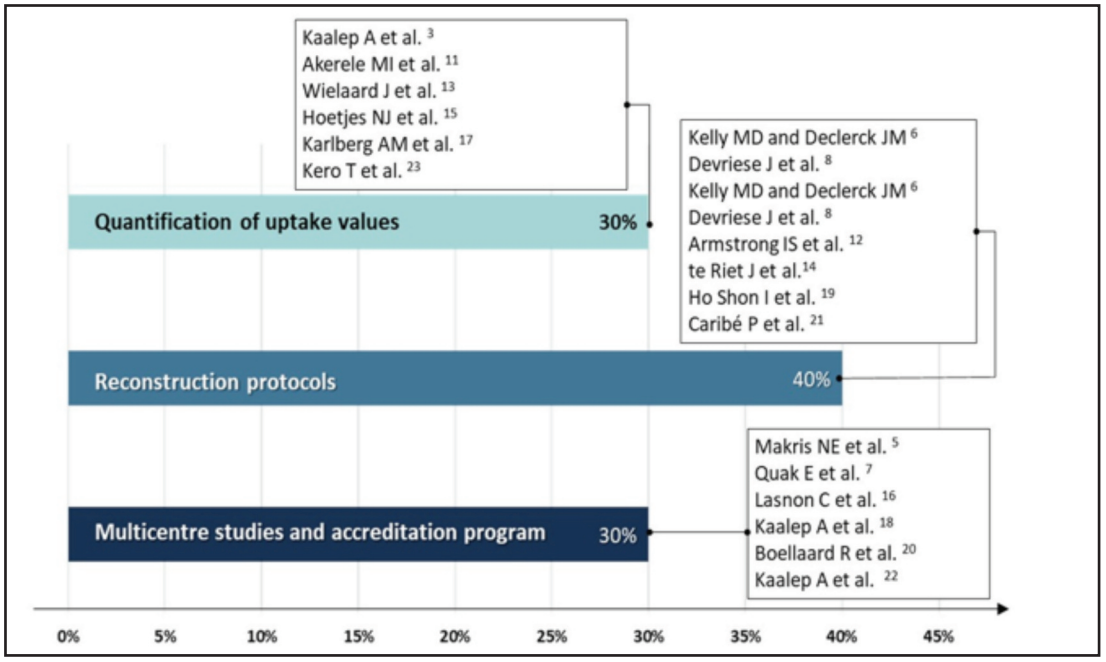


Fig. 3: The result of phantoms application in PET/CT studies.

tumour (sphere of phantom). As the spheres of the phantom are fixed, the results obtained by Hoetjes NJ et al. were consistent during the simulation. This proves that phantoms are acceptable and reliable for image quantification and interpretation.

Besides, Karlberg AM et al. showed the phantom's role in a study of the quantitative comparison of PET system performance between PET/CT scanner and PET/MR scanner.[17] They utilised NEMA phantom to run the performance of PET image quality. They reported that the mean of hot lesions for the systems was relatively similar, and the SUV quantifications showed an acceptable agreement for these two PET systems.[17] The experiment also used the same homogeneous NEMA phantom for the systems to prevent the heterogeneous activity ratios between spheres and background.

Therefore, many clinical researchers used phantoms as a combined experiment since the quantification of the phantoms is consistent. For a new tracer apart of ^{18}F-FDG, the phantom is capable of showing the reliability for image quantification. For example, Kero T et al. used 11C, and Wielaard J et al. used 68Ga in their studies.[13,23] Kero T et al. studied the SUV quantification 11C in cardiac clinical research.[23]

Wielaard J et al. presented the method to ascertain a noise level principle for ^{68}Ga-PSMA PET/CT imaging studies using NEMA phantom. The phantom results validated the significance of the recommended activity procedure on image quantification using 15 retrospective PET/CT patient scans.[13] The method determined that the minimum injected dose for clinical ^{68}Ga-PSMA imaging studies was acceptable for quantification and was also reliable for image interpretation of PET/CT. This shows the need for extensive phantom studies since there are no clear guidelines for tracer of ^{68}Ga-PSMA in PET/CT system, unlike the ^{18}F-FDG PET/CT guidelines that were well established. The ^{18}F-FDG PET/CT guideline will be discussed in the subtopic of "application of phantoms in multicentre and accreditation programme ".

Application of phantoms in multicentre studies and accreditation programme

Phantoms are routinely used as the first step of quality control in examining PET/CT systems to test a PET/CT system's performance, directly impacting the clinical outcome. Recently, many researchers put phantoms as the main subject in multicentre studies. The outcomes of their studies were managed to recommend the specific guidelines, standards, and accreditation programmes. Most researchers have growingly conducted multicentre studies of PET image quality of SUVs between phantoms and patient data throughout the past decade.[32]

Makris NE et al. confirmed that PET/CT institutions need to harmonise the scanners among the various institutions when conducting multicentre trials. All three of the different phantoms (NEMA, ACR and Anthropomorphic) tested in their study were suitable for harmonising various PET/CT scanners' quantitative performance in the Netherlands.[5] However, the article did not state any accreditation programme but suggested that PET/CT be harmonised, as the study was a multicentre PET/CT study.

In France, two articles demonstrated the multicentre study in tumour PET imaging. Quak E et al. used NEMA phantom and oncology patients (n=517) to perform optimal lesion detection.[7] They harmonised the quantification of lesion detection from a single PET acquisition and processed the data set. Meanwhile, Lasnon C et al. utilised NEMA phantom to study the harmonising SUVs in multicentre trials.[16] They followed the accreditation programme of the European Association of Nuclear Medicine (EANM) guidelines to harmonise quantitative values.

In addition, EANM proposed a new guideline of harmonisation of PET/CT – Version 2.0 as the "EANM Procedure Guidelines for Tumour Imaging: Version 2.0". This guideline's aim is to update the guideline version 1.0 that published in 2010.[20,33] Accuracy and precision are also crucial as ^{18}F-FDG PET/CT is utilised to assess tumour response and diagnosis, prognosis, and staging either in phantom or clinical purposes.[20] Besides, Kaalep A et al. reported that the EANM Research Ltd. (EARL) had collected more than 2500 phantom datasets from 200 PET/CT scanners and including 150 worldwide imaging institutions under project of PET/CT accreditation programme.[18] Under the EARL initiative, the EANM has been running an ^{18}F-FDG PET/CT accreditation programme to harmonise the quantitative PET/CT performance and assist multicentre nuclear medicine and research.

We find that the guidelines for ^{18}F-FDG tracer of PET/CT tumour imaging are well established. They utilised various phantom studies to establish the guidelines. However, the rest of the tracers, such as ^{68}Ga, ^{11}C and ^{89}Zr, demonstrate multicentre studies but still have not established any recommended guideline.[13,22,23] Kaalep A et al. suggested the immediate action to develop an applicable cross-calibration and accreditation programme to guide multicentre 89Zr quantitative studies.[22] Perhaps we will see the clarity of the accreditation programme in the future for the new PET/CT tracers rather than the ^{18}F-FDG only.

Application of phantoms in Reconstruction Protocol Studies of PET/CT

The reliance of PET/CT accuracy on quantitative uptake values obtained with different reconstruction protocols uses phantoms with uniformed geometry and activity preparation, presenting an acceptable estimation of clinical morphology and activity administration.[33] The phantoms are a beneficial benchmark of PET/CT scanner performance, integrating the effects of detector resolution, scanner sensitivity, the accuracy of the various corrections performed, and the reconstruction parameters used. For example, the number of iterations and subsets and post-filter smoothing reconstruction parameter. Reconstruction set-ups should be determined for PET/CT scanner capable of producing resolution recovery coefficients within the specified bounds to construct the images.[6] Therefore, Kelly M and Declerck J introduced a new reference standardised uptake value (SUVref) procedure in PET/CT reconstruction protocol by experimenting the NEMA phantom to minimise PET/CT scanner hardware variability.[6]

Devriese J et al. also introduced a new reconstruction protocol approach based on Q. Clear reconstruction criteria by utilising the NEMA phantom experiment and according to Belgian law. They are mainly used for the treatment response assessment. The harmonisation reconstruction protocol should be used as the reconstruction protocol for lesion detection using resolution recovery technique. This reconstruction protocol complies with the EARL programme specifications.[8]

Rogasch J et al. and Armstrong I et al. studied the impact of TOF and PSF reconstruction, focusing on phantom and lung lesions, respectively. Rogasch J et al. used cylindrical phantom to study the TOF and PSF reconstruction outcomes of uptake values.[9] While, Armstrong I et al. implemented the TOF and PSF reconstruction of NEMA phantom data with patient data. However, some minor variations in clinical data existed when TOF was implemented, which was not seen in phantom data, which required further study.[12]

Recently, a reconstruction algorithm method of Bayesian penalised-likelihood (BPL) has been established to produce images with enhanced signal-to-noise ratio and minimised noise compared to standard ordered subsets expectation maximisation (OSEM) algorithm.[35] Ishimori T et al. investigated the effect of Bayesian penalised-likelihood (BPL) PET reconstruction condition on quantitative parameters in FDG-PET/CT.[35] Nevertheless, it does not clarify how the condition of BPL reconstruction influences the quantification of clinical PET/CT.

Then, Riet J et al. presented PET/CT performance evaluation using Bayesian penalised-likelihood (BPL) reconstruction to study the realistic clinical conditions using phantoms-patient-based experiments. They used two types of the phantom which are NEMA and a modified micro hollow sphere phantom. They found that the BPL algorithm's performance is remarkable to the common OSEM-PSF algorithm in tiny and small-scale lesion detectability.[14]

Caribé P et al. also focused on the BPL reconstruction study of NEMA phantom and Belgian patient retrospective studies.[21] They found that BPL can minimise the noise compared to the OSEM reconstruction method. In 2020, Texte E et al. used different phantoms such as ACR phantom to evaluate PET/CT BPL reconstruction's effect in small lesion detectability using low contrast. Interestingly, they used ^{18}F-FDG and hypoxia tracers such as ^{18}F-MIZO and ^{18}F-FAZA.10 Interestingly, the result was consistent with all other previous papers studying this reconstruction algorithm.

Next, the reviewed article states the reconstruction protocol related to CT reconstruction instead of PET reconstruction. Ho Shon I et al. demonstrated the effect of CT reconstruction algorithms and acquisition parameters on attenuation correction for PET reconstruction. They undertook an anthropomorphic torso phantom study to assess CT acquisition parameters' impact with the lower dose. CT iterative reconstruction enhances image quality with lower exposures. Very low dose CT exposures are possible for accurate PET attenuation correction. They suggested that the scanner and reconstruction-specific validation should be employed prior very low dose CT for PET.[19]

Limitations and Future Approach

This systematic review has some limitations. One of the drawbacks is the search strategy, which included only original English-language research papers published between 2010 and 2020. As a result, the probability of sample bias exists in this study. We also acknowledge that the data presented was just a small data collection regarding the type of phantoms used for PET/CT imaging studies (n=20). However, this small data enables us to clarify the need of a system phantom for PET/CT imaging.

Furthermore, another factor limiting the efficiency of the systematic review is that it only coveres the homogenous phantoms study rather than the heterogeneous phantoms study. The heterogeneous phantoms can be used to simulate real clinical conditions. This could be the suggestion of improvement for future research.

Overall, the heterogeneity of phantoms in the literature suggests that further approach research is needed to study the recent advances in phantom development, especially in 3D phantom printing. We believe that further investigation will lead to better consistency in quantitative PET/CT hybrid imaging for diagnostic application. Perhaps, this can be implemented for other applications such as theranostics and dosimetry applications. In the future, this could pave the way for modern medical physics and molecular imaging field studies in Malaysia.

CONCLUSION

This review contributes an overview of the need for a system phantom for quantitative hybrid PET/CT scans despite limited guidance and literature about this topic. According to this review, the phantom study has a pivotal role in hybrid nuclear imaging of PET/CT either in technical aspects of the scanners (such as data acquisition and reconstruction protocol) or clinical characteristics of patients. This study identified the need for phantoms used within quantitative hybrid PET/CT scans, especially for quantification, optimisation, harmonisation and standardisation of PET/CT scanners.

Besides, the necessity to identify the suitable system phantoms to utilise within PET/CT scans by considering the continuous research and keeping ongoing to study a new phantom development. Researchers are encouraged to adopt efforts on phantom quantitative validation, including verification with clinical data of patients. Perhaps, researchers could take into consideration the continuing development of new phantom technologies innovation in the future.

CONFLICT OF INTEREST

We inform that there is no conflict of interest regarding the publication of this paper. We reported that all articles met the stated relevant requirements, but if some papers are missing, this was not intended and we would like to apologise to any authors.

ACKNOWLEDGEMENTS

We acknowledged the Ministry of Higher Education Malaysia via the Fundamental Research Grant Scheme (FRGS/1/2018/STG02/UPM/02/7) for support of publication expenses.

REFERENCES

1. Bertolini V, Trojani V, Bertolini M. CT protocol optimisation in PET/CT: what we learn from a systematic review. Eur J Nucl Med Mol Imaging 2021; 48(1): 1–2.
2. Hristova I, Boellaard R, Galette P, Shankar LK, Liu Y, Stroobants S, et al. Guidelines for quality control of PET/CT scans in a multicenter clinical study. EJNMMI Phys 2017; 4(1): 23.
3. Kaalep A, Burggraaff CN, Pieplenbosch S, Verwer EE, Sera T, Zijlstra J, et al. Quantitative implications of the updated EARL 2019 PET–CT performance standards. EJNMMI Phys 2019; 6(1): 1-16.
4. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg 2010; 8(5):336-41.
5. Makris NE, Huisman MC, Kinahan PE, Lammertsma AA, Boellaard R. Evaluation of strategies towards harmonisation of FDG PET/CT studies in multicentre trials: comparison of scanner validation phantoms and data analysis procedures. Eur J Nucl Med Mol Imaging 2013; 40(10): 1507-15.
6. Kelly MD, Declerck JM. SUVref: Reducing reconstruction-dependent variation in PET SUV. EJNMMI Res 2011; 1(1): 1-11.
7. Quak E, Le Roux PY, Hofman MS, Robin P, Bourhis D, Callahan J, et al. Harmonising FDG PET quantification while maintaining optimal lesion detection: prospective multicentre validation in 517 oncology patients. Eur J Nucl Med Mol Imaging 2015; 42(13): 2072-82.
8. Devriese J, Beels L, Maes A, Van de Wiele C, Pottel H. Impact of PET reconstruction protocols on quantification of lesions that fulfil the PERCIST lesion inclusion criteria. EJNMMI Phys 2018; 5(1): 1-13.
9. Rogasch JMM, Hofheinz F, Lougovski A, Furth C, Ruf J, Großer OS, et al. The influence of different signal-to-background ratios on spatial resolution and F18-FDG-PET quantification using point spread function and time-of-flight reconstruction. EJNMMI Phys 2014 ; 1(1): 1-16.
10. Texte E, Gouel P, Thureau S, Lequesne J, Barres B, Edet-Sanson A, et al. Impact of the Bayesian penalised likelihood algorithm (Q.Clear®) in comparison with the OSEM reconstruction on low contrast PET hypoxic images. EJNMMI Phys 2020; 7(1): 1-15
11. Akerele MI, Wadhwa P, Silva-Rodriguez J, Hallett W, Tsoumpas C. Validation of the physiological background correction method for the suppression of the spill-in effect near highly radioactive regions in positron emission tomography. EJNMMI Phys 2018 ; 5(1): 1-21
12. Armstrong IS, Kelly MD, Williams HA, Matthews JC. Impact of point spread function modelling and time of flight on FDG uptake measurements in lung lesions using alternative filtering strategies. Arch der Math 2014; 1(1): 1-18.
13. Wielaard J, Habraken JBA, Brinks P, Lavalaye J, Boellaard R. Optimization of injected 68Ga-PSMA activity based on list-mode phantom data and clinical validation. EJNMMI Phys 2020; 7(1): 1-12.
14. te Riet J, Rijnsdorp S, Roef MJ, Arends AJ. Evaluation of a Bayesian penalised likelihood reconstruction algorithm for low-count clinical 18F-FDG PET/CT. EJNMMI Phys 2019; 6(1): 1-14.
15. Hoetjes NJ, Van Velden FHP, Hoekstra OS, Hoekstra CJ, Krak NC, Lammertsma AA, et al. Partial volume correction strategies for quantitative FDG PET in oncology. Eur J Nucl Med Mol Imaging 2010; 37(9): 1679-87.
16. Lasnon C, Desmonts C, Quak E, Gervais R, Do P, Dubos-Arvis C, et al. Harmonising SUVs in multicentre trials when using different generation PET systems: Prospective validation in non-small cell lung cancer patients. Eur J Nucl Med Mol Imaging 2013; 40(7): 985-96.
17. Karlberg AM, Sæther O, Eikenes L, Goa PE. Quantitative comparison of PET performance—siemens biograph mCT and mMR. EJNMMI Phys 2016; 3(1):1-14.
18. Kaalep A, Sera T, Oyen W, Krause BJ, Chiti A, Liu Y, et al. EANM/EARL FDG-PET/CT accreditation - summary results from the first 200 accredited imaging systems. Eur J Nucl Med Mol Imaging 2018; 45(3): 412-2.
19. Ho Shon I, Reece C, Hennessy T, Horsfield M, McBride B. Influence of X-ray computed tomography (CT) exposure and reconstruction parameters on positron emission tomography (PET) quantitation. EJNMMI Phys 2020; 7(1): 1-16.

20. Boellaard R, Delgado-Bolton R, Oyen WJG, Giammarile F, Tatsch K, Eschner W, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. Eur J Nucl Med Mol Imaging 2015; 42(2): 328-54.
21. Caribé PRRV, Koole M, D'Asseler Y, Van Den Broeck B, Vandenberghe S. Noise reduction using a Bayesian penalised-likelihood reconstruction algorithm on a time-of-flight PET-CT scanner. EJNMMI Phys. 2019; 6(1): 1-14.
22. Kaalep A, Huisman M, Sera T, Vugts D, Boellaard R. Feasibility of PET/CT system performance harmonisation for quantitative multicentre 89Zr studies. EJNMMI Phys 2018; 5(1): 1-8.
23. Kero T, Sörensen J, Antoni G, Wilking H, Carlson K, Vedin O, et al. Quantification of 11C-PIB kinetics in cardiac amyloidosis. J Nucl Cardiol 2020; 27(3): 774-84.
24. Data Spectrum. NEMA IEC PET Body Phantom Set [cited 8 Feb 2021]. Available from: https://www.spect.com/our-products/acceptance-testing/nema-iec-pet-body phantom-set/
25. Biodex. PET Phantom - NEMA 2012; [cited 8 Feb 2021]. Available from: https://www.biodex.com/nuclear-medicine/products/pet-positron-emission-tomography/pet-phantoms/pet-phantom-nema-2012
26. Biodex. Flangeless Deluxe PET and SPECT Phantoms; [cited 8 Feb 2021]. Available from: https://www.biodex.com/nuclear-medicine/products/pet-positron-emission-tomography/pet-phantoms/flangeless-deluxe-pet-and-sp
27. Supertech, Inc. Jaszczak Flangeless Deluxe PhantomTM; [cited 8 Feb 2021]. Available from:https://www.supertechxray.com/NuclearMedicine/JaszczakACRSPECTPhantoms.php
28. Mawlawi OR, Jordan DW, Halama JR, Schmidtlein CR, Wooten WW. PET/CT acceptance testing and quality assurance: The report of AAPM task group 126. American Association of Physicists in Medicine 2019; 48(2): 1-40.
29. Biodex. Anthropomorphic SPECT Phantom; [cited 8 Feb 2021]. Available from: https://www.biodex.com/nuclear-medicine/products/phantoms/anthropomorphic-spect-phantom
30. Data Spectrum. Micro Hollow Sphere Phantom; [cited 8 Feb 2021]. Available from: https://www.spect.com/our-products/preclinical/micro-hollow-sphere-phantom/
31. Ziai P, Hayeri MR, Salei A, Salavati A, Houshmand S, Alavi A, et al. Role of optimal quantification of FDG PET imaging in the clinical practice of radiology. Radiographics 2016; 36(2): 481–96.
32. Hanafi MH, Noor NM, Rana S, Saad FFA. Standardisation Techniques of Independent PET/CT Modalities Utilising PET SUVmax as a Potential Conversion Marker. Transylvanian Rev 2017; 1(7).
33. Boellaard R, O'Doherty MJ, Weber WA, Mottaghy FM, Lonsdale MN, Stroobants SG, et al. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. Eur J Nucl Med Mol Imaging 2010; 37(1): 181.
34. Gnesin S, Kieffer C, Zeimpekis K, Papazyan JP, Guignard R, Prior JO, et al. Phantom-based image quality assessment of clinical 18F-FDG protocols in digital PET/CT and comparison to conventional PMT-based PET/CT. EJNMMI Phys 2020; 7(1): 1–16.
35. Ishimori T, Nakamoto Y, Miyake K, Saga T, Koyasu S, Togashi K. Effect of Bayesian penalized-likelihood PET reconstruction condition on quantitation. J Nucl Med 2019; 60(supplement 1): 1601