ORIGINAL ARTICLE

Preclinical SPECT and PET: Joint EANM and ESMI procedure guideline for implementing an efficient quality control programme

Christian Vanhove[1] · Michel Koole[2] · Pedro Fragoso Costa[3] · Margret Schottelius[4] · Julia Mannheim[5] · Claudia Kuntner[6] · Geoff Warnock[7,8] · Wendy McDougald[9,10] · Adriana Tavares[9,11] · Monique Bernsen[12]

Abstract
The aim of this guideline is to provide recommendations for the implementation of an effective and efficient quality control (QC) programme for SPECT and PET systems in a preclinical imaging lab. These recommendations aim to strengthen the translational power of preclinical imaging results obtained using preclinical SPECT and PET. As for clinical imaging, reliability, reproducibility, and repeatability are essential when groups of animals are used in a longitudinal imaging experiment. The larger the variability of the imaging endpoint, the more animals are needed to be able to observe statistically significant differences between groups. Therefore, preclinical imaging requires quality control procedures to maintain reliability, reproducibility, and repeatability of imaging procedures, and to ensure the accuracy and precision of SPECT and PET quantification. While the Physics Committee of the European Association of Nuclear Medicine (EANM) has already published excellent procedure guidelines for Routine Quality Control Recommendations for Nuclear Medicine Instrumentation that also includes procedures for small animal PET systems, and important steps have already been made concerning preclinical quality control aspects, this new guideline provides a review and update of these previous guidelines such that guidelines are also adapted to new technological developments.

Keywords Preclinical · SPECT · PET · Quality control · Guideline

Preamble

The European Society for Molecular Imaging (ESMI) represents the scientific community involved in multidisciplinary molecular imaging science. This includes basic, preclinical, translational, and clinical research on imaging technologies, methodologies, biomarkers and probes, modelling and data analysis. The European Association of Nuclear Medicine (EANM) is a professional non-profit medical association that facilitates communication worldwide between individuals pursuing clinical and research excellence in nuclear medicine. The EANM was founded in 1985.

ESMI and EANM have members who are scientists specialising in the research and practice of nuclear medicine. The ESMI and EANM periodically define new guidelines to help advance the science of nuclear medicine. Existing practice guidelines will be reviewed for revision or renewal, as appropriate, on their fifth anniversary or sooner, if indicated.

Each practice guideline, representing a policy statement by the ESMI/EANM, has undergone a thorough consensus process in which it has been subjected to extensive review. The ESMI and EANM recognize that the safe and effective use of diagnostic nuclear medicine imaging requires specific training, skills, and techniques, as described in each document. Reproduction or modification of the published practice guideline by those entities not providing these services is not authorized.

These guidelines are an educational tool designed to assist practitioners. They are not inflexible rules or requirements of practice and are not intended, nor should they be used, to establish a legal standard. For these reasons and those set forth below, both the ESMI and the EANM caution against the use of these guidelines in litigation in which the decisions of a practitioner are called into question.

The ultimate judgment regarding the propriety of any specific procedure or course of action must be made by the principal investigators in light of all the circumstances presented. Thus, there is no implication that an approach differing from the guidelines, standing alone, is below the

---

Extended author information available on the last page of the article

standard. To the contrary, a conscientious practitioner may responsibly adopt a course of action different from that set forth in the guidelines when, in the reasonable judgment of the practitioner, such course of action is indicated limitations of available resources, or advances in knowledge or technology subsequent to publication of the guidelines.

Introduction

Preclinical studies using laboratory animal models are a crucial step towards understanding the underlying mechanism of disease through observations of responses to interventions and physiological and/or environmental changes at tissue, cell or molecular level. They are an essential component of biomedical research to explore, develop, and validate new drugs, biomarkers, or therapies. Nowadays, a wide range of animal models that mimic different human pathologies, such as oncological, cardiological, neurological and inflammatory diseases, are available [1, 2] with 10.8 million animals used for medical research in 2018 in the EU alone as demonstrated by European Union (EU) statistics [3]. As such, the preclinical field bridges the more fundamental scientific discoveries at molecular/cellular level with the clinical implementation of new diagnostics and/or therapeutics. With the continuous growth of pharmaceutical and biotechnology companies, this importance of translating scientific discoveries into practical clinical applications is expected to further increase in the future.

In preclinical settings, traditional investigative approaches include histology and organ (tissue) sampling. While such techniques have provided valuable insights into the biochemistry, cellular and molecular mechanisms of diseases, they present some important limitations [4]: (1) In vitro methods do not visualize the entire organism at once and over time; (2) These methods require removal of cells/tissue from their native environment and sample processing. The use of fixatives perturbs the natural micro-environment such that results cannot be considered fully representative of the true pathophysiological conditions; (3) Because of the invasiveness of these methods, animals generally need to be sacrificed for each investigative time point, which not only makes longitudinal measurement impossible but also greatly increases the number of animals in each of the experimental and control subgroups [5, 6]. To overcome these limitations, preclinical imaging has become a valuable research tool for longitudinal in vivo studies of laboratory animal models of disease and for animal models validation using diagnostic procedures applied in clinical practice. Using molecular imaging techniques, such as PET and SPECT, radioligands and radiotracers can be used for example to determine receptor availability and rate of biological processes in vivo. The added value of imaging resides in its non-invasive nature to acquire quantitative anatomical, functional, and molecular information in intact organisms, with the unique opportunity to repeatedly and precisely follow-up the disease course and therapeutic responses. This also relates to the fact that the number of animals required for a particular study can be considerably reduced and further refined by using preclinical imaging techniques, which complies with the ethical 3R policies (reduction, refinement and replacement) devised by Russell and Burch [7, 8]. During longitudinal imaging studies, the animals are not sacrificed and they can thus each serve as their own control, such that biological variability is greatly alleviated at the benefit of statistical power [5]. The implementation of preclinical in vivo imaging technologies in the biomedical research arena has ever since revolutionized the way that research studies are conducted to answer biological and/or clinical questions, thereby accelerating the transfer of laboratory discoveries into clinical practice. With the rapid advent of preclinical imaging, we assume that a large part of the animals that are used in the EU for medical research [3] would also be part of preclinical imaging experiments.

Generally speaking, preclinical imaging systems are modified versions of their clinical equivalent [4]. They essentially rely on the same underlying physical principles, but there are major differences to account for when scaling-down a clinical to a preclinical imager. The volume (size) of the subject to be imaged and the spatial resolution needed to detect anatomical or functional changes are profoundly different; a 25 g mouse is about 3000 times smaller in volume than a 75 kg human. This implies that preclinical imaging systems require 10 times better spatial resolution compared to clinical imaging systems to attain similar discriminatory capabilities. To obtain this better spatial resolution, preclinical PET systems are designed as scaled down versions of clinical systems, use very similar image formation techniques as clinical PET, and spatial resolutions of 1 mm can be obtained. On the other hand, most preclinical SPECT systems use pinhole collimators that are mainly used for (planar) thyroid imaging in clinical practice. Similar to radionuclide thyroid imaging, the magnification on the projection images due to the pinhole geometry can dramatically reduce information loss as a result of the intrinsic spatial resolution of the gamma detector when an animal is brought in close proximity to the pinhole opening, which can result in sub-mm spatial resolutions [6]. In addition to dimensional scaling, there are also functional aspects to be considered, such as differences in heart and respiratory cycle between rodents and humans, which calls for a better temporal resolution. For example, a cardiac gated-SPECT or gated-PET acquisition of a mouse with an organ definition and a temporal resolution that is comparable to that of a human scan, requires a sub-millimetre spatial resolution and a ten-fold increase in frame rate [9].

Compared to clinical systems, small animal systems are also different in the way they are used. Small animal imaging has seen a significant growth during the last decade, and dedicated systems for imaging rats and mice have been built and are now commercially available. Based on a recent preclinical imaging market report (https://www.researchreportsworld.com/enquiry/request-sample/21076013) the global preclinical imaging market, including PET and SPECT but also MRI, CT, US and optical imaging, is expected to further grow with 6.5% from 2022 until 2028. As of today, there is a large variety of systems offered by different vendors. In general, these systems are installed in a research facility (university, pharmaceutical or biotechnology company) and they might not be used daily, as opposed to patient examinations in a clinical environment. The systems are often used on a more irregular basis, usually the systems will be used quite intensively during research studies where numerous animals will be scanned with the same acquisition protocol, sequentially or by scanning multiple animals simultaneously. Another important difference to the clinical situation is that while clinical imaging systems are supported by a medical physics expert or medical physicist, this support is often lacking in a preclinical setting. One of the key activities of the medical physicist and medical physics expert, as described by the European Federation of Organizations for Medical Physics (EFOMP) [10], is: 'The specification, management and supervision of associated quality assurance/control programmes.' This implies that in the clinical situation, it is clear who is responsible for monitoring the QC programme, while these responsibilities are less well defined in the preclinical situation. Other differences with clinical systems are the wider range of studied contrast agents, radiotracers, and isotopes. These are usually novel imaging probes that need to be evaluated preclinically before they can be translated into the clinic. Finally, it has been recognized that, more often than not, preclinical imaging standards fall shorter than those expected and used in the clinical setting [11–13].

In a recent survey of the Standard study group of the European Society for Molecular Imaging (ESMI), where one of the major aims was to gather knowledge on the current state-of-the-art of preclinical imaging quality control (QC) procedures used at different sites and for different instruments (*i.e.*, SPECT, PET, CT, MRI, US and optical imaging), 47% of survey participants ($n = 71/151$) didn't have or didn't know about QC guidelines at their institute [14]. Of all the survey participants, 46% use preclinical PET ($n = 61/132$) and 20% ($n = 27/132$) are using preclinical SPECT. QC might appear less important in preclinical imaging as it is not aimed at directly producing diagnostic information, however, its translational potential certainly support clinical research making QC procedures essential. Furthermore, scientific results based on sub-optimal preclinical imaging is undesirable, especially in view of the mandatory integrity and quality of scientific data. Some examples of the consequences of poor QC can be found in the work of McDougald WA and Mannheim JG [15]. But, efforts to define and implement standardized imaging procedures to deliver reliable, reproducible, repeatable, and translatable preclinical imaging data sets has been limited so far [11, 15, 16]. Meanwhile, the ethical rule of the 3Rs exists since 1959 [7] and all reputable funding bodies and scientific journals require adherence to this ethical principles, but there are no such requirements for QC procedures. Because the major strength of preclinical imaging is longitudinal follow-up of animals, preclinical imaging devices should be operational as reliable measurement tools with required regular quality control procedures [12]. Therefore, to ensure that studies using preclinical imaging technologies have the highest level of scientific integrity and adhere to the 3Rs by not having to extend the number of animals due to non-reliable system performance, a community-led consensus for the implementation of QC programme for preclinical imaging is warranted.

Goal

The aim of this guideline is to provide recommendations for the implementation of a QC programme for PET and SPECT systems in preclinical imaging labs and assist users with QC procedures.

We are convinced that this is still essential, especially for preclinical SPECT. Evidence from the above mentioned ESMI survey indicated that for preclinical SPECT only 36% ($n = 9/25$), 40% ($n = 10/25$) and 44% ($n = 11/25$) of the survey participants uses QC procedures such as photopeak drift, uniformity testing and collimator checking, respectively [14]. For preclinical PET, QC procedures were clearly more common, with 75% ($n = 42/56$) of users regularly performing scanner quality control (34% daily ($n = 19/56$), 20% weekly ($n = 11/56$) and 21% monthly ($n = 12/56$)). Preclinical SPECT users also indicated that 85% of the experiments involves quantitative SPECT imaging. Therefore, we are still convinced that it is necessary to create procedure guidelines for imaging laboratories in order to set up an efficient QC programme.

Criteria for the QC programme

To ensure that the imaging results obtained from a laboratory animal imaging experiment are reliable, reproducible, and repeatable several factors have to be taken into account (Fig. 1) [15].
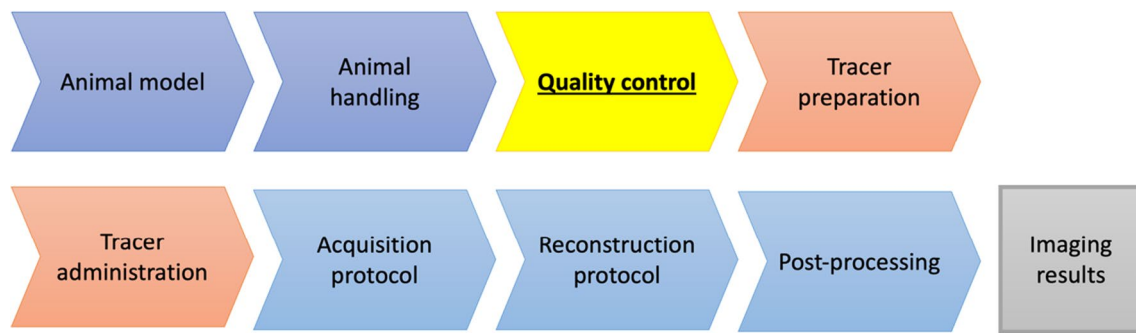
Fig. 1 Chronological order of steps conducted during a preclinical imaging experiment and that have an influence on the final imaging results

The animal models used, how the animal is treated in between imaging sessions, the route of injection of the radiotracer, how the animal is prepared right before the imaging study (*i.e.*, anaesthesia used), how the animal is treated during imaging (*i.e.*, heating, level and length of anaesthesia), and how the animal is monitored during imaging (*i.e.*, cardiac and/or respiratory triggering) will have an effect on the final imaging results. The technical performance of the equipment under routine conditions, which can be assessed using QC procedures, will also affect the final imaging results. Similar to the clinic, the isotope used, tracer selection, tracer production and labelling are key components that influence final imaging results in PET and SPECT. Finally, the acquisition and reconstruction protocol settings of the imaging system, together with any other image processing steps, obviously also will have an important effect on the final image. QC should be one of the first steps in the chain in laboratory animal imaging experiments (see Fig. 1) and is absolutely necessary, because even if all other aspects are standardised a technically failing system will result in incorrect imaging data.

Criterium 1: In this guideline, we will focus on QC procedures of PET and SPECT instrumentation to ensure that the QC metrics under routine conditions are stable over time.

In 2008, the National Electrical Manufacturers Association (NEMA) has published a document dealing with performance measurements of small-animal PET scanners (NEMA NU 4–2008) [17]. The goal of the NEMA NU 4–2008 is to enable comparison of the performance of different preclinical PET systems over a wide range of technologies and geometries used. Most research prototype PET systems and virtually all commercially available preclinical PET scanners have published performance evaluations based on the NEMA standard [18–20]. However, the NEMA NU-4 recommendations do not take into account several important parameters, such as performance with different isotopes,

quantitative heterogeneity across the field-of-view, and the use of different reconstruction methods. But, as shown in the work of Disselhorst et al. [21], the NEMA NU-4 image quality phantom can be used to compare image quality parameters when using different PET isotopes and reconstruction methods. The NEMA NU 4–2008 standard was devised almost 16 years ago, and recently Hallen et al. examined if the NEMA standard still meets its goals to enable a fair comparison of PET systems taken into account newer technological developments and paradigm shifts [22]. Unfortunately, no such consensus document exists (yet) for small-animal SPECT scanners. The reason for this is probably associated to the larger variety in equipment hardware and acquisition parameters settings in SPECT compared to PET. Variety in equipment hardware includes planar versus cylindrical pinhole collimators, static versus rotating collimators, and number of pinholes used. For SPECT acquisitions, isotopes with different photon energy are used, in contrast to the fixed value of 511 keV in PET. This also requires the selection of an appropriate collimator (low energy *versus* high energy). In addition, the radius of rotation can also be set as an acquisition parameter in some systems. All these items could have led to a more difficult consensus in creating a NEMA standard for preclinical SPECT compared to preclinical PET. Although, the NEMA standard is (1) an essential benchmark in the development of new PET systems to determine peak performance, (2) an important tool for the purchase decisions of potential buyers, and (3) by most vendors used as the gold standard for acceptance testing to assess whether their equipment meets all specifications; they are not primarily meant for recommendations with regard to routine QC procedures. Specifically, a QC programme should include test frequencies and action levels, so that test results can be compared to predefined values. These aspects were not the objective of the NEMA consensus, but as stated above elements of NEMA, such as the NEMA-NU4 image quality phantom, have been proven useful to compare image quality parameters under different imaging conditions [21].

Criterium 2: This guideline will use elements of the NEMA NU 4–2008 recommendations to evaluate the stability of image quality metrics.

This guideline will not focus on acceptance testing, but this guidelines will further build on the good practice QC paper of Osborne et al. [23], the standardization paper of Mannheim et al. [16], and the opinion paper of McDougald et al. [15]. The first paper considered a good benefit–cost ratio by suggesting QC tests that were designed with the knowledge that not all preclinical imaging laboratories have access to specialized phantoms, and the tests can be performed with items commonly found in many laboratories, but the suggested QC tests provide sufficient performance feedback to evaluate the status of an imaging platform. The second and third paper proposed to extend the QC measurements with the use of phantoms (*i.e.*, NEMA NU-4 image quality phantom), so that QC metrics can be used whose values can be required to fall within a range of reference values (*e.g.*, recovery coefficients). We would support the inclusion of phantom measurement in this guideline as it would open the way towards harmonization of preclinical imaging. It should also be noted that by using the NEMA NU-4 image quality phantom, three important QC metrics can be derived from one measurement: image uniformity, quantitative accuracy, and recovery coefficient. These parameters are directly measured on reconstructed images and from a practical point of view one could decide to only monitor such image-derived parameters in a QC programme. But one might run the risk that if there is a slow deviation of PET/SPECT detector(s) performance, it might not be immediately visible in the 'imaging domain' [15].

Criterium 3: In this guideline we will use both image-derived QC metrics and more fundamental measurements in 'detector space' (*e.g.*, photopeak position).

Preclinical imaging systems are often installed in a research environment where available resources and personnel are limited, which can contribute to the limited ability to implement a QC programme. Therefore, to ensure that as many preclinical imaging laboratories as possible implement a QC programme and thus to guarantee proper daily functioning of preclinical PET and SPECT equipment, QC tests should be straightforward, not time-consuming and the number of tests to be performed should be limited. On the other hand, preclinical imaging is an important translational tool that requires imaging results to be reliable, reproducible, repeatable, and accurate, thus demanding for objective QC metrics to evaluate the performance of the equipment used. However, the costs and benefits of a QC programme should be well balanced. Examples of costs are the costs of phantoms (with prices

ranging for a few hundred to thousand euros), radioactive sources (also with prices ranging for a few hundred to thousand euros), staff time (5–10% FTE), downtime of the equipment, and radiation burden to the personnel. The benefit is related to the probability of detecting any degradation of the imaging system and avoiding its consequences. The consequence, or the 'cost', of not implementing a QC programme can lead to the late detection of total detector failure, resulting in unusable PET or SPECT data that may lead in the need to use a larger number of animals (and thus not complying to the ethical rule of the 3Rs). Next to total system failure, a system can also drift over time and the impact of system drift will be much more difficult to detect without a QC programme. For example, a drift in the photopeak of Tc-99 m of 5% can result in a decrease of the signal-to-noise from 12.1 to 9.1 (unpublished data using uniform phantom filled with 7.7 MBq/ml). This will lead to larger variability in the imaging data (*i.e.*, noise increases from 8.3% to 11.0%) and as a consequence, based on power analysis, the minimum required sample size would have increase from 8 to 12 animals to observe a 20% difference between two groups of animals (and thus also not complying to the ethical rule of the 3Rs). Another example of larger variability in the imaging data as a results of a suboptimal preclinical PET system is illustrated in the work of McDougald WA and Mannheim JG [15].

Criterium 4: This guideline will provide recommendations for the implementation of an effective and efficient QC programme for preclinical imaging laboratories using SPECT and/or PET to check reliability under routine conditions with good benefit-to-cost ratio. In term of the cost–benefit trade-off, the more frequently QC tests recommended in this guideline are of very short duration and do not require dedicated phantoms.

Most performance metrics can be easily adapted from PET to SPECT. QC metrics that can be used for PET and SPECT are, for instance, image uniformity, quantitative accuracy, and recovery coefficients. On the other hand, the large variety of SPECT isotopes with different energies, and accordingly the different collimators recommended for these SPECT-isotopes, calls for special attention in SPECT. In principle, this requires that performance parameters should be measured for each collimator-isotope pair separately. However, in order to balance costs and benefits, QC procedures for the most commonly used collimator-isotope pair should be performed more frequently than other combinations.

Criterium 5: This guideline will use the same performance metrics for PET and SPECT.

General recommendations

This guideline covers the types of QC tests to be performed, the purpose of these QC tests, the description of the test protocols and the suggested frequencies. In this section, we will provide some general recommendations regarding implementing a QC programme and how action levels can be set.

The first step that must be performed when starting a QC programme is to define reliable baseline values for the QC tests results. These baseline values will be the reference for future QC results and are crucial to evaluate the stability of the equipment over time. Acceptance testing will play a crucial role here and ensures that the PET and SPECT systems meet the specified performance criteria and quality standards when a new system is installed or after a major system upgrade and before they are put into routine use. Because preclinical SPECT and PET devices are expensive instruments, preclinical imaging facilities could include the need for acceptance testing in the tendering process when purchasing new equipment. For PET, these acceptance tests can include the NEMA NU 4–2008 performance measurements and for SPECT these tests can include sensitivity, spatial resolution, energy resolution, count-rate versus activity, and uniformity across the field-of-view [24]. Once the instrument has been accepted for routine use, it is important to execute all QC tests initially in order to obtain appropriate baseline values for future cross-referencing of QC test results. If baseline values after acceptance testing are not available, baseline values may be obtained after a preventive maintenance of the imaging system. Based on the results of the above-mentioned ESMI survey, more than 70% of the SPECT or PET users have such preventive maintenance services performed at least once a year [14].

QC test results need to be recorded in a logbook. This is crucial for identifying performance changes over time. Ideally, this logbook is a digital or electronic record because it creates easy access and search options to the data, anywhere and anytime on nearly any device. Furthermore, a digital record offers the flexibility to implement (automated) computational processing tasks, such as the regular generation of a graph plotting the QC results over time (together with the reference value); thereby visualizing trends in performance that may be indicative to predict or reveal upcoming system failures problems. The logbook should include the serial number, type and vendor of the preclinical PET or SPECT system, together with the dates, nature and results of the QC tests.

An electronic logbook can also help to define action levels. Action levels should be set to maintain variations in system performance, as identified by QC test results, within certain limits. While QC test procedures and suggested frequencies can be described in a very precise way, this does not apply to the way action levels can be determined. In general, it is very difficult to define absolute values for action levels ab initio. Instead, action levels should be determined locally on the basis of experience, and considering the manufacturer's recommendations, with the cost–benefit aspect in mind. Therefore, action levels should be set so that, on the one hand, system degradations have not reached a stage where they can be detected in the preclinical routine imaging results. On the other hand, action levels should not be too strict, causing long downtimes for the preclinical imaging equipment for repairs, readjustments, and calibration.

Any follow-up actions taken following unsatisfactory QC test results should also be recorded in the logbook, as this may assist troubleshooting when similar problems (re-) occur.

Last but not least, as mentioned in the introduction section, an important difference with the clinical situation is that support of a medical physics expert or medical physicist is often missing in the preclinical field. Monitoring the QC programme is the responsibility of the medical physics expert or medical physicist, and these responsibilities should also be defined in preclinical imaging laboratories. The responsible person for the QC programme in the preclinical situation will be often the facility manager or the staff that is responsible for the day-to-day activities in the preclinical imaging lab. The responsibility for a QC programme should be clearly assigned and agreed to ensure that the QC programme can be properly followed up.

Action levels

To assist preclinical imaging labs to set action levels, we suggest the concept of the cumulative sum procedure. The cumulative sum procedure is a statistical process control method that employs a statistical method for monitoring change detection on sequential data [25], such as for example QC results logged over period of time. The cumulative sum procedure involves plotting:

$$S_t = max(0, S_{t-1} + Z_t - \omega) \tag{1}$$

at the tth observation, where $S_0 = 0$ and Z_t is the sample weight assigned to the tth observation. ω is an optional and tuneable parameter that can be used to adjust the sensitivity of change detection. A larger ω-value will make the cumulative sum procedure less sensitive to change.

When using quality control results, we recommend applying the following procedure:

- Calculate the expected mean μ_x and standard deviation σ_x of a set of quality control results x_t obtained after system installation or preventive maintenance and considered to be appropriate for baseline or reference values.
- Calculate the Z-score for each quality control results:

$$Z_t = (x_t - \mu_x)/\sigma_x \qquad (2)$$

Table 1 Example of the use of the cumulative sum procedure using 20 observations. The second column (x_t) shows the logged values of a quality control metric, the third column (Z_t) the Z-score of this metric, the fourth column the cumulative sum St with $\omega = 0$, and the fifth column the cumulative sum S_t with $\omega = 1$

Observation	x_t	Z_t	S_t with $\omega = 0$	S_t with $\omega = 1$
1	3.70	0.67	0	0
2	3.34	-0.53	0	0
3	3.27	-0.78	0	0
4	3.72	0.72	0.72	0
5	3.33	-0.57	0.15	0
6	3.27	-0.77	0	0
7	3.31	-0.63	0	0
8	3.15	-1.17	0	0
9	3.12	-1.25	0	0
10	3.30	-0.68	0	0
11	3.75	0.83	0.83	0
12	3.36	-0.47	0.37	0
13	3.68	0.60	0.96	0
14	3.32	-0.60	0.36	0
15	3.26	-0.80	0	0
16	3.28	-0.73	0	0
17	3.61	0.37	0.37	0
18	3.81	1.03	1.40	0.03
19	3.94	1.47	2.87	0.50
20	4.02	1.73	4.60	1.23

- Calculate the cumulative sum for each Z-score using Eq. 1. We recommend using an ω-value of 1.
- When S_t becomes larger than 0, an action is expected.

An example is given in Table 1 and illustrated in Fig. 2. The example shows 20 observations of a QC measure (x_t) with a mean of 3.5% and a standard deviation of 0.3% as baseline values. From the Z_t column it can be observed that the Z-score of these 20 observations never deviate by more than 2 standard deviations. The S_t column (with $\omega = 0$) shows that the cumulative sum sometimes becomes greater than zero when QC results become larger than the baseline value. But the cumulative sum becomes zero again after (a number of) observations that are smaller than the baseline value. However, from the 17th observation QC results are systematically above the baseline value that might indicate an upcoming system failure. As a result, the cumulative sum value (with $\omega = 0$) becomes increasingly larger as can be observed in the fourth column of Table 1. By tuning ω to a value of 1, the cumulative sum becomes greater than zero only from the 17th observation, which could be a threshold to take action. As already mentioned, the Z-scores in Table 1 never deviate by more than 2 standard deviations, indicating that action levels only based on high deviations from baseline will not necessarily detect changes over time.

Recommended PET QC tests

This section presents the QC tests recommended for preclinical PET. The purpose of the QC tests, the tools and equipment required, the description of the test protocols, and the suggested frequencies are described. All the recommended tests for preclinical PET are also summarised in Table 2.
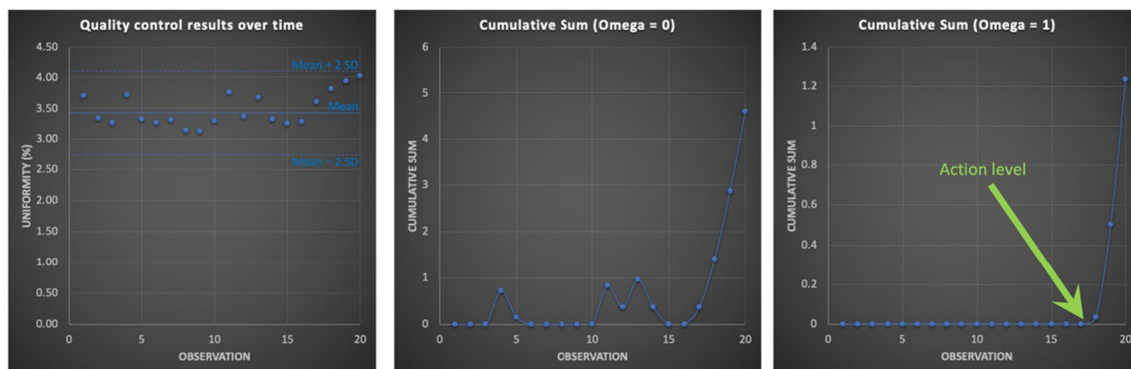

Fig. 2 Cumulative sum procedure. The figure on the left shows 20 observations of quality control results with a mean value of 3.5% and a standard deviation (SD) of 0.3%. The cumulative sum chart with $\omega = 0$ is shown in the middle. The figure on the right shows the cumulative sum chart with $\omega = 1$, where a value > 0 could be a threshold to take action

Table 2 Overview of recommended QC tests for preclinical PET. The purpose of the QC tests, the recommended frequency, the source required, and the suggested radioactivity are listed

Routine test	Purpose	Frequency	Source	Activity
Physical inspection	To identify mechanical and other defects of gantry and handling system that may cause system failure	Daily (recommended) Certainly on a scan day (minimal)	Not applicable	Not applicable
Detector check	To check that all detectors/blocks are working	Daily (recommended) Certainly on a scan day (minimal)	Point source	0.37–0.74 MBq
Image uniformity	To estimate voxel-value uniformity by imaging a uniformly filled object	Monthly	Cylindrical phantom	3.7—10 MBq
Quantitative accuracy	Determine that calibration constant is constant over time	Monthly	Cylindrical phantom with known activity concentration	3.7—10 MBq
Recovery coefficients	Determine that spatial resolution is constant over time using the recovery coefficients concept	Yearly	NEMA NU 4–2008 image quality phantom (or alternative)	3.7 MBq
PET/CT alignment	Determine possible misalignments between PET and CT component	Yearly	Cylindrical phantom, capillary tubes or zeolite beads	3.7—10 MBq

Physical inspection

Purpose

A QC programme is aimed at detecting changes that occur so slowly that they would not be noticed in everyday use, however, it should be mentioned that a QC programme can never replace the normal attentiveness of the operator in observing equipment problems that are immediately obvious. Therefore, a simple physical inspection of the imaging system (gantry and handling system) for mechanical or other defects might already prevent system failure.

Tools and equipment required

None.

Procedure

Physical inspection of the imaging system (gantry and handling system) for mechanical or other defects that may cause system failure. Some components that are recommended to be checked are bed connections, anaesthesia supply, animal heating, and the monitoring of respiratory and/or cardiovascular signals.

Frequency

Daily (recommended), or at least during a scan day prior to scanning (minimal).

Detector check

Purpose

Check that all detector elements of the preclinical PET systems are working properly.

Tools and equipment required

Mostly a sealed radioactive source of Na-22 or Ge-68, or a point source of an easily available positron emitter is required. A partially drawn syringe, a capillary tube or microcentrifuge tubes can be used as point source proxies. In some PET systems with lutetium-based (Lu-176) scintillators, the intrinsic Lu-176 radiation emitted from the scintillators is used to check the detectors and here no additional isotopes are required.

Procedure

Most vendors provide procedures to check that the detector elements are working properly and/or to check the 511 keV

energy peak. We recommend using these vendor procedures to visually evaluate that all detector elements are working properly and that there is no photopeak drift. Generally, these procedures require a point source with low radioactivity (0.37—0.74 MBq) that have to be placed in the centre of the field of view. After a short acquisition a detector efficiency map is shown. All detector elements should be functioning properly. Screenshots of these detector efficiency maps can be stored.

We recommend that photopeak drift should be < 10%.

Frequency

Daily (recommended), or at least before use during a scan day (minimal).

Image uniformity

Purpose

This test measures that the voxel-values in a reconstructed image are uniformly distributed when acquiring a homogeneous radioactivity distribution.

Tools and equipment required

A cylindrical hollow phantom with a volume of roughly 30 ml and a diameter corresponding to the diameter of a mouse (± 30 mm). Centrifuge tubes or liquid scintillator vials can be used as refillable cylindrical phantom. Furthermore, an easily available positron emitter, such as F-18, is required to fill this phantom. In preclinical laboratories where a positron emitter other than F-18 is routinely used (*e.g.*, Ga-68, Cu-64, Zr-89, etc.), this isotope can be used to evaluate uniformity [21]. Also, commercially available uniform cylindrical phantoms that use long-living isotopes, such as Ge-68 and Na-22, can be used for evaluating uniformity.

Procedure

When a refillable phantom is used, the phantom has to be filled with a solution of water and the isotope. The radioactivity of the positron emitter used to fill the phantom should ideally replicate the routine conditions for most imaging experiments that are performed in the preclinical laboratory (typically 3.7—10 MBq). One should take care that the solution is uniformly distributed in the cylinder and that air bubbles are not present. Ensure that the chemical form of the isotope is compatible with the phantom material to avoid adsorption of radionuclides on the walls of phantoms [26]. The cylinder should also be sealed tightly. A PET-scan of the cylindrical phantom has to be acquired with the

phantom positioned at the centre of the scanner and using an acquisition time that is similar to routinely used acquisition times in the lab (for example 10—20 min). The acquired data can be reconstructed using the reconstruction algorithm and reconstruction parameters that are used under routine condition. Although attenuation and scatter are relatively low in mice, we recommend correcting for attenuation, and preferably also for scatter, during image reconstruction. In case of multi-centre studies, it is highly recommended to use a standardized reconstruction protocol as mentioned by McDougald et al. [11]. After reconstruction, a cylindrical region-of-interest has to be drawn centrally within the phantom with a diameter and a height equal to 75% of the phantom diameter and height, respectively. The standard deviation divided by the mean (*i.e.*, coefficient of variation) of the voxel-values within the region of interest has to be logged, should be stable over time, and we recommend this value should be < 10%.

Frequency

Monthly. After any detector relevant hardware replacement (*i.e.*, detectors blocks or cooling systems).

Quantitative accuracy

Purpose

This test evaluates the ability of the PET-system to correctly measure radioactivity concentrations. To correctly measure radioactivity concentrations, the calibration constant should be constant over time (as its name suggests). The calibration constant is used to correlate the radioactivity measured using a dose calibrator and radioactivity concentrations measured on a reconstructed PET-image.

Tools and equipment required

A refillable cylindrical phantom with a known volume (of roughly 30 ml) and a diameter corresponding to the diameter of a mouse (± 30 mm). Centrifuge tubes, syringes or liquid scintillator vials can be used as refillable cylindrical phantom. Furthermore, an easily available positron emitter, such as F-18, is required to fill this phantom. In preclinical laboratories where a positron emitter other than F-18 is routinely used (*e.g.*, Ga-68, Cu-64, Zr-89, etc.), this isotope can be used to evaluate quantitative accuracy. A quality-controlled dose calibrator is required to accurately measure the radioactivity of the positron emitter. Recommended procedures for dose calibrator quality control programmes have previously been described in the EANM guideline [27]. Also, commercially available uniform cylindrical phantoms that use long-living isotopes, such as Ge-68 and Na-22, can also be

used for evaluating quantitative accuracy when the activity concentration in these phantoms is known.

Procedure

When a refillable phantom is used, the phantom has to be filled with a solution of water and the isotope. By calculating the phantom volume using inner diameter and height, by measuring the volume during filling, or by weighing the phantom empty and full the correct volume within the phantom can be obtained. The radioactivity of the positron emitter used to fill the phantom should ideally replicate the routine conditions for most imaging experiments that are performed in the preclinical laboratory (for examples, 3.7—10 MBq) and should be accurately measured in a syringe using the dose calibrator available in the preclinical imaging lab, including background activity subtraction. The residual radioactivity in the syringe after filling the phantom should also be measured using this dose calibrator. One should take care that the solution is uniformly distributed in the cylinder and that air bubbles are not present. Ensure that the chemical form of the isotope is compatible with the phantom material to avoid adsorption of radionuclides on the walls of phantoms [26]. The cylinder should also be sealed tightly. By knowing the volume of the phantom, the radioactivity of the full syringe and the radioactivity remaining in the syringe, the radioactivity concentration, expressed in MBq/ml, in the phantom is known. A PET-scan of the cylindrical phantom has to be acquired with the phantom positioned at the centre of the scanner and using an acquisition time that is similar to routine used acquisition times in the lab (for example, 10—20 min). The acquired data can be reconstructed using the reconstruction algorithm and reconstruction parameters that are used under routine condition. We highly recommend correcting for attenuation, and preferably also for scatter, during image reconstruction. In case of multi-centre studies, it is highly recommended to use a standardized reconstruction protocol [11]. After reconstruction a cylindrical region-of-interest has to be drawn centrally within the phantom with a diameter and a height equal to 75% of the phantom diameter and height, respectively. By dividing the known radioactivity concentration in the phantom with the average of the voxel-values within the region of interest, the so-called Q-factor can be calculated. This Q-factor has to be logged and should be stable over time for each individual isotope that is frequently used (< 10% variation over time). Also note that with the procedure described here, both image accuracy and image uniformity can be evaluated together. Therefore, to improve the efficiency (cost–benefit) of the QC procedure, we recommend that image uniformity and quantitative accuracy can be performed in one phantom preparation.

Frequency

Monthly. After any detector relevant hardware replacement (*i.e.*, detectors blocks or cooling systems).

Recovery coefficients

Purpose

To evaluate that the spatial resolution of a PET-scanner is stable over time. In this guideline the recovery coefficient (RC) concept will be used as its direct relation with spatial resolution has been previously described [28].

Tools and equipment required

To determine recovery coefficients on preclinical PET-systems the fillable NEMA NU-4 2008 image quality phantom (66 mm long, 33.5 mm diameter) should be used. An alternative design of this phantom is also acceptable. The NEMA image quality phantom consists of 3 different regions to evaluate image uniformity, RCs and spill-over [17, 18, 20, 22]. RCs can be determined using the 5 hot rods with different diameters (5, 4, 3, 2, and 1 mm) and the region to evaluate image uniformity. An easily available positron emitter, such as F-18, is required to fill this phantom. In preclinical laboratories where a positron emitter other than F-18 is routinely used (*e.g.*, Ga-68, Cu-64, Zr-89, etc.), this isotope can be used to (indirectly) evaluate spatial resolution.

Procedure

The NEMA-protocol can be used for this QC test where the phantom is filled with an activity of 3.7 MBq. One should take care that the solution is uniformly distributed in the cylinder and that air bubbles are not present. Ensure that the chemical form of the isotope is compatible with the phantom material to avoid adsorption of radionuclides on the walls of phantoms [26]. A 20-min PET-scan of the image quality phantom has to be acquired using with the phantom positioned at the centre of the scanner. The acquired data can be reconstructed using the reconstruction algorithm and reconstruction parameters that are used under routine condition. We recommend correcting for attenuation, and preferably also for scatter, during image reconstruction. In case of multi-centre studies, it is highly recommended to use a standardized reconstruction protocol as mentioned by McDougald et al. [11]. To calculate the RCs, a four-step procedure is required: (1) the reconstructed transverse image slices along the central 10 mm of the hot rods have to be averaged, (2) circular regions of interest, with diameters twice the diameter of the hot rod, have to be drawn over the six hot rods on this averaged slice, (3) a cylindrical region

of interest having a length of 10 mm and a 25 mm diameter (*i.e.*, covering 75% of the central uniform section) has to be drawn centrally in the uniform region of the phantom, and (4) the RCs can be calculated as the maximum voxel-value in the circular regions of interest (after averaging transverse image slices along the central 10 mm of the hot rods section of the phantom) divided by the mean voxel-value measured in the cylindrical region-of-interest. The RCs should be close to 1.0 for rods with a diameter larger that 3 times the expected system spatial resolution, close to 0.75 for rods with a diameter twice the expected system spatial resolution, and less than 0.5 for rods with a diameter equal to or smaller than the expected system spatial resolution [18–21, 29]. Because the RCs are defined as the maximum values in a circular region of interest around the hot rods, divided by the mean activity in a volume of interest over the uniform region, it is important to note that this definition of the RCs does not measure the mean recovery in the hot rods but a combination of recovery and variance [22]. This means that RCs can be overestimated in case of high variance/noise and thus that RCs can be > 1 in case good recovery but high variance/noise. RCs have to be logged and should be stable over time. Note that with the procedure described here, image uniformity, image accuracy and RCs can be evaluated using only one acquisition when the volume of the phantom is known, and the radioactivity used to fill the phantom is accurately measured using a dose calibrator. So, when this phantom is available in a preclinical lab, this might improve the efficiency of the QC procedure.

Frequency

Yearly. This procedure can also be followed when the image quality of different PET scanners at different preclinical imaging sites needs to be harmonised.

PET/CT alignment

Purpose

To determine possible misalignments between PET and CT components. Misalignments between PET and CT components can occur due to drifts in mechanical aspects of the gantry, such as accuracy of bed positioning and slight movement of components. Misalignment between PET and CT images will impact quantitative accuracy when CT images are used to correct for attenuation and/or scatter, and when CT images are used as anatomical reference to draw regions of interest for PET image quantification. Important to note here is that a QC procedure is also required for CT imaging and that this procedure confirms that the CT device is functioning properly [15, 16, 23].

Tools and equipment required

Most vendors provide phantoms to test for possible misalignment between their multimodal components. However, as alternative a refillable cylindrical hollow phantom, centrifuge tubes or liquid scintillator vials can also be used. Other options are capillary tubes (minimum 2) or zeolite beads (minimum 3) asymmetrically placed on a cylindrical object. Furthermore, an easily available positron emitter, such as F-18, is required to fill the phantom, vial, tubes or to be soaked by the beads.

Procedure

When vendors provide phantoms and procedures to determine any misalignment between their multimodal components, we recommend using these vendor procedures to visually determine any possible misalignments between PET and CT. When these procedures are not available, a PET and a CT scan of the filled phantom, vial, tubes or infused zeolite beads has to be acquired using the same animal bed and without repositioning the phantom, vial, tubes or zeolite beads. The acquired PET/CT data can be reconstructed using the reconstruction algorithm and reconstruction parameters that are used under routine condition. Using the laboratory's multimodal image viewer possible misregistration between the PET and CT needs to be determined.

Frequency

Yearly. After any replacement of detector relevant hardware or animal bed.

Recommended SPECT QC tests

This section presents the QC tests recommended for preclinical SPECT. The purpose of the QC tests, the tools and equipment required, the description of the test protocols, and the suggested frequencies are described. All the recommended tests for preclinical SPECT are also summarised in Table 3.

Physical inspection

Purpose

For SPECT, it is certainly very important to mention that a QC programme can never replace the normal attentiveness of the operator in noticing problems with equipment that are immediately obvious. Therefore, a simple physical inspection of the imaging system for mechanical or other defects might already prevent system failure.

Table 3 Overview of recommended QC tests for preclinical SPECT. The purpose of the QC tests, the recommended frequency, the source required, and the suggested radioactivity are listed

Routine test	Purpose	Frequency	Source	Activity	Comments
Physical inspection	Check for collimator damage and collimator mountings to detector head	Daily (recommended) Certainly on a scan day (minimal)	Not applicable	Not applicable	Inspect for mechanical and other defects that may compromise safety of animals. If collimator damage is detected or suspected, a uniformity test should be performed
Detector check and photopeak drift	To check that gamma-detectors are working	Daily (recommended) Certainly on a scan day (minimal)	Point source	± 1 MBq	When using a single-photon emitter that is not frequently used, detector check and photopeak drift should be performed prior to scanning
Image uniformity	To estimate voxel-value uniformity by imaging a uniformly filled object	Monthly, using the collimator that is most frequently used in daily practice, and a single-photon emitter that is easily available	Cylindrical phantom	10—20 MBq	When using a collimator-isotope combinations that is not frequently used, image uniformity should be checked prior to scanning
Quantitative accuracy	Determine that calibration constant is constant over time	Monthly, using the collimator that is most frequently used in daily practice, and a single-photon emitter that is easily available	Cylindrical phantom with known activity concentration	10—20 MBq	When using a collimator-isotope combinations that is not frequently used, quantitative accuracy should be checked prior to scanning
Recovery coefficients	Determine that spatial resolution is constant over time using the recovery coefficients concept	Yearly, using the collimator that is most frequently used in daily practice, and a single-photon emitter that is easily available	NEMA NU 4–2008 image quality phantom (or alternative)	37 MBq	Can be used to harmonize SPECT acquisitions and reconstruction
SPECT/CT alignment	Determine possible misalignments between PET and CT component	Yearly, using the collimator that is most frequently used in daily practice, and a single-photon emitter that is easily available	Cylindrical phantom, capillary tubes or zeolite beads	10—20 MBq	When using a collimator-isotope combinations that is not frequently used, SPECT/CT alignment should be checked prior to scanning

Tools and equipment required

None.

Procedure

Performing a physical inspection for any damage to the collimator(s) and collimator mounts on the gamma detector heads can prevent poor image quality and avoid safety issues for animals. If collimator damage is detected or suspected, a uniformity test should be immediately performed (see details below). Other components that are recommended to be checked are bed connections, anaesthesia supply, animal heating, and the monitoring of respiratory and/or cardiovascular signals.

Frequency

Daily (recommended), or at least during a scan day prior to scanning (minimal).

Detector check and photopeak drift

Purpose

Check that all detectors are working properly.

Tools and equipment required

Mostly a sealed radioactive source (such as Co-57) or a point source of an easily available single-photon emitter is required. A partially draw syringe, a capillary tube or microcentrifuge tubes can be used as point sources.

Procedure.

Most vendor provide procedures to check that the gamma detectors are working properly and to check photopeak drift. Some procedures are performing this functional test with collimators (*i.e.*, extrinsic uniformity), but generally a procedure without collimators is used (*i.e.*, intrinsic uniformity). We recommend using these vendor procedures to visually evaluate that all detector elements are working properly and that there is no photopeak drift. Generally, these procedures require a point source with low radioactivity (1 MBq) that have to be placed in the centre of the field of view without using a collimator (intrinsic uniformity). In the presence of a collimator, most procedures use a sheet source (*e.g.*, Co-57) to evaluate extrinsic uniformity. After a short acquisition detector flood maps are shown, together with the position and the width of the photopeak. Detector flood maps should be uniform. Screenshots of detector flood maps can be stored

and photopeak positions can be logged. Photopeak positions should be stable over time and should not differ more than 10% from the actual photopeak position of the isotope.

Frequency

Daily (recommended), or at least before use during a scan day (minimal). When using a single-photon emitter that is not frequently used, detector check and photopeak drift should also be performed prior to scanning.

Image uniformity

Purpose

This test measures that the voxel-values in a reconstructed image are uniformly distributed when acquiring a homogeneous radioactivity distribution.

Tools and equipment required

A cylindrical hollow phantom with a volume of roughly 30 ml and a diameter corresponding to the diameter of a mouse (± 30 mm). Centrifuge tubes or liquid scintillator vials can be used as refillable cylindrical phantom. Furthermore, an easily available single-photon emitter, such as Tc-99 m, is required to fill this phantom. In preclinical laboratories where a single-photon emitter other than Tc-99 m is routinely used (*e.g.*, I-123, In-111, etc.), this isotope can be used to evaluate uniformity.

Procedure

The phantom has to be filled with a solution of water and the isotope. The radioactivity of the single-photon emitter used to fill the phantom should ideally replicate the routine conditions for most imaging experiments that are performed in the preclinical laboratory (for example, 10—20 MBq). One should take care that the solution is uniformly distributed in the cylinder and that air bubbles are kept to a minimum. Ensure that the chemical form of the isotope is compatible with the phantom material to avoid adsorption of radionuclides on the walls of phantoms [26]. The cylinder should also be sealed tightly. A SPECT-scan of the cylindrical phantom has to be acquired using an acquisition time that is similar to routine used acquisition times in the lab (for example, 20—30 min). SPECT acquisition and reconstruction protocols are based on collimator selection, photopeak definition, and corrections. The collimator to be used should be the collimator most commonly used in the lab for SPECT acquisitions. Ideally, image uniformity should be evaluated using every possible collimator-isotope combination used in a preclinical lab. However, this does not

fit within the cost–benefit principle and therefore we recommend evaluating image uniformity using the collimator that is most commonly used in daily practice, in combination with a single-photon emitter that is easily available in the preclinical imaging lab. The acquired data can be reconstructed using the reconstruction algorithm and reconstruction parameters that are used under routine condition using a correct photopeak(s) selection. Although attenuation and scatter are relatively low in mice, we still recommend correcting for attenuation, and preferably also for scatter, during image reconstruction. In case of multi-centre studies, it is highly recommended to use a standardized reconstruction protocol [11]. After reconstruction a cylindrical region-of-interest has to be drawn centrally within the phantom with a diameter and a height equal to 75% of the phantom diameter and height, respectively. The standard deviation divided by the mean (*i.e.*, coefficient of variation) of the voxel-values within the region of interest has to be logged, should be stable over time, and < 10%.

Frequency

Monthly, or at least during a scan day prior to scanning when using a collimator-isotope combinations that is not frequently used. After any detector relevant hardware replacement.

Quantitative accuracy

Purpose

This test evaluates the ability of the SPECT-system to correctly measure radioactivity concentrations. To correctly measure radioactivity concentrations, the calibration constant should be constant over time (as its name suggests). The calibration constant is used to correlate the radioactivity measured using a dose calibrator and radioactivity concentrations measured on a reconstructed SPECT-image. Based on the ESMI-survey, the majority (85%) of SPECT-experiments involve quantitative imaging [14].

Tools and equipment required

A refillable cylindrical phantom with a known volume (of roughly 30 ml) and a diameter corresponding to the diameter of a mouse (± 30 mm). Centrifuge tubes or liquid scintillator vials can be used as refillable cylindrical phantom. Furthermore, an easily available single-photon emitter, such as Tc-99 m, is required to fill this phantom. In preclinical laboratories where a single-photon emitter other than Tc-99 m is routinely used (*e.g.*, I-123, In-111, etc.), this isotope should be used to evaluate quantitative accuracy. A quality-controlled dose calibrator is required to accurately measure the

radioactivity of the single-photon emitter. Recommended procedures for dose calibrator quality control programmes have previously been described in the EANM guideline [27].

Procedure

The phantom has to be filled with a solution of water and the isotope. By calculating the phantom volume using inner diameter and height, by measuring the volume during filling, or by weighing the phantom empty and full the correct volume within the phantom can be obtained. The radioactivity of the single-photon emitter used to fill the phantom should ideally replicate the routine conditions for most imaging experiments that are performed in the preclinical laboratory (for example 10—20 MBq) and should be accurately measured in a syringe using the dose calibrator available in the preclinical imaging lab, including background activity subtraction. The residual radioactivity in the syringe after filling the phantom should also be measured using this dose calibrator. One should take care that the solution is uniformly distributed in the cylinder and that air bubbles are kept to a minimum. Ensure that the chemical form of the isotope is compatible with the phantom material to avoid adsorption of radionuclides on the walls of phantoms [26]. The cylinder should also be sealed tightly. By knowing the volume of the phantom, the radioactivity of the full syringe and the radioactivity remaining in the syringe, the radioactivity concentration, expressed in MBq/ml, in the phantom is known. A SPECT-scan of the cylindrical phantom has to be acquired using an acquisition time that is similar to routinely used acquisition times in the lab (for example 20—30 min). SPECT acquisition and reconstruction protocols are based on collimator selection, photopeak definition, and corrections. The collimator to be used should be the collimator most commonly used in the lab for SPECT acquisitions. Ideally, image uniformity should be evaluated using every possible collimator-isotope combination used in a preclinical lab. However, this does not fit within the cost–benefit principle and therefore we recommend to evaluate quantitative accuracy using the collimator that is most commonly used in daily practice, in combination with a single-photon emitter that is easily available in the preclinical imaging lab. The acquired data can be reconstructed using the reconstruction algorithm and reconstruction parameters that are used under routine condition using a correct photopeak(s) selection. We highly recommend correcting for attenuation, and preferably also for scatter, during image reconstruction. In case of multi-centre studies, it is highly recommended to use a standardized reconstruction protocol as mentioned by McDougald et al. [11]. After reconstruction, a cylindrical region-of-interest has to be drawn centrally within the phantom with a diameter and a height equal to 75% of the phantom diameter and height, respectively. By dividing the

known radioactivity concentration in the phantom with the average of the voxel-values within the region of interest, the so-called Q-factor can be calculated. This Q-factor has to be logged and should be stable over time for each individual isotope that is frequently used (< 10% variation over time). Also note that with the procedure described here, both image accuracy and image uniformity can be evaluated together. Therefore, to improve the efficiency (cost–benefit) of the QC procedure, we recommend that image uniformity and quantitative accuracy can be performed in one phantom preparation.

Frequency

Monthly, or at least during a scan day prior to scanning when using a collimator-isotope combinations that is not frequently used. After any detector relevant hardware replacement.

Recovery coefficients

Purpose

To evaluate that the spatial resolution of a SPECT scanner is stable over time. In this guideline the recovery coefficient (RC) concept will be used as its direct relation with spatial resolution has been previously described for PET scanners [28]. We believe the same procedure can be used for preclinical SPECT.

Tools and equipment required

Therefore, to determine recovery coefficients on preclinical SPECT-systems the NEMA NU-4 2008 phantom, or an alternative design of this phantom, can be used. An easily available single-photon emitter, such as Tc-99 m, is required to fill this phantom. In preclinical laboratories where a single-photon emitter other than Tc-99 m is routinely used (*e.g.*, I-123, In-111, etc.), this isotope can be used to evaluate spatial resolution.

Procedure

We recommend filling the phantom with an activity of 37 MBq. One should take care that the solution is uniformly distributed in the cylinder and that air bubbles are kept to a minimum. Ensure that the chemical form of the isotope is compatible with the phantom material to avoid adsorption of radionuclides on the walls of phantoms [26]. A 20-min SPECT-scan should be acquired of the image quality phantom and the acquired data should be reconstructed using the reconstruction algorithm and reconstruction parameters that are used under routine condition using a correct photopeak(s) selection. We recommend correcting for attenuation, and preferably also for scatter, during image reconstruction. In case of multi-centre studies, it is highly recommended to use a standardized reconstruction protocol [11]. The collimator to be used should be the collimator most commonly used in the lab for SPECT acquisitions. Ideally, RCs have to be evaluated using every possible collimator-isotope combination used in a preclinical lab. However, this does not fit within the cost–benefit principle and therefore we recommend determining the RCs using the collimator that is most commonly used in daily practice, in combination with a single-photon emitter that is easily available in the preclinical imaging lab. To calculate the RCs, a four-step procedure is required: (1) the reconstructed transverse image slices along the central 10 mm of the hot rods have to be averaged, (2) circular regions of interest, with diameters twice the diameter of the hot rod, have to be drawn over the six hot rods on this averaged slice, (3) a cylindrical region of interest having a length of 10 mm and a 25 mm diameter (*i.e.*, covering 75% of the central uniform section) has to be drawn centrally in the uniform region of the phantom, and (4) the RCs can be calculated as the maximum voxel-value in the circular regions of interest (after averaging transverse image slices along the central 10 mm of the hot rods section of the phantom) divided by the mean voxel-value measured in the cylindrical region-of-interest. RCs have to be logged and should be stable over time. The RCs should be close to 1.0 for rods with a diameter larger that 3 times the expected system spatial resolution, close to 0.75 for rods with a diameter twice the expected system spatial resolution, and less than 0.5 for rods with a diameter equal to or smaller than the expected system spatial resolution [18–21, 29]. Because the RCs are defined as the maximum values in a circular region of interest around the hot rods, divided by the mean activity in a volume of interest over the uniform region, it is important to note that this definition of the RCs does not measure the mean recovery in the hot rods but a combination of recovery and variance [22]. This means that RCs can be overestimated in case of high variance/noise and thus that RCs can be > 1 in case good recovery but high variance/noise. RCs have to be logged and should be stable over time. Note that with the procedure described here, image uniformity, image accuracy and RCs can be evaluated together when the volume of the phantom is known, and the radioactivity used to fill the phantom is accurately measured using a dose calibrator. So, when this phantom is available in a preclinical lab, this might improve the efficiency of the QC procedure.

Frequency

Yearly. This procedure can also be followed when the image quality of different SPECT scanners at different preclinical imaging sites needs to be harmonised.

SPECT/CT alignment

Purpose

To determine possible misalignments between SPECT and CT components. Misalignments between SPECT and CT components can occur due to drifts in mechanical aspects of the gantry, such as accuracy of bed positioning and slight movement of components. Misalignment between SPECT and CT images will impact quantitative accuracy when CT images are used to correct for attenuation and when CT images are used as anatomical reference to draw region of interest for SPECT image quantification. Important to note here is that a QC procedure is also required for CT imaging and that this procedure confirms that the CT device is functioning properly [15, 16, 23].

Tools and equipment required

Most vendors provide phantoms to test for possible misalignment between their multimodal components. However, a refillable cylindrical hollow phantom, centrifuge tubes or liquid scintillator vials can also be used. Other options are capillary tubes (minimum 2) or zeolite beads (minimum 3) asymmetrically placed on a cylindrical object. Furthermore, an easily available positron emitter, such as Tc-99 m, is required to fill the phantom, vial, tubes or to be soaked by the beads.

Procedure

When vendors provide phantoms and procedures to determine any misalignment between their multimodal components, we recommend using these vendor procedures to visually determine any possible misalignments between SPECT and CT. When these procedures are not available, a SPECT and a CT scan of the filled phantom, vial, tubes or infused zeolite beads has to be acquired using the same animal bed and without repositioning the phantom, vial, tubes or zeolite beads. The acquired SPECT/CT data can be reconstructed using the reconstruction algorithm and reconstruction parameters that are used under routine condition. Using the laboratory's multimodal image viewer possible misregistration between the PET and CT needs to be determined.

Frequency

Yearly, or at least during a scan day prior to scanning when using a collimator-isotope combinations that is not frequently used. After any replacement of detector relevant hardware (including collimator) or animal bed.

Summary and conclusions

As for clinical applications, preclinical imaging requires QC procedures to maintain reliability of the results and to ensure accuracy of SPECT and PET quantification. However, it has been recognized that, more often than not, preclinical imaging standards fall shorter than those expected and used in a clinical setting. Although we are convinced that implementing a QC program might be challenging, we encourage small animal laboratories to follow the basic guidelines outlined here since it will certainly increase the confidence of the results obtained from imaging experiments towards results obtained by other research groups and when contracting imaging services. In this guideline, we have tried to balance the cost and benefits because we are aware that preclinical imaging systems are often installed in a research environment where available resources and personnel are limited. To ensure that as many preclinical imaging laboratories as possible implement a QC programme and thus to guarantee proper daily functioning of preclinical PET and SPECT equipment, the QC tests presented here should be straightforward and not too time-consuming.

This guideline also has some limitations that are imported to mention. The consequence of QC programme with a good cost–benefit balance is that only the most commonly used isotopes will be used for performing the QC procedures. However, it should be noted that when using other isotopes, there is a potential risk for lower-quality data and inaccurate quantification. For example, the use of positron emitters with a longer positron travel will result in poorer spatial resolution in the reconstructed images, and to ensure accurate quantification the PET and SPECT scanner should be correctly calibrated for a particular isotope to obtain the radionuclide-specific calibration factor that will allow correct conversion from measured counts/voxel into radionuclide concentration measurements. Furthermore, in some preclinical imaging labs multiple-animal beds are used and it is important to consider the impact on image quality and quantitative accuracy of scanning multiple animals in the same field of view. The presence of more than one concentrated source of radioactivity may negatively affect attenuation, increase the singles and randoms rates, increase the number of scatter events, increase the detector and system dead time, and reduce resolution and sensitivity as subjects are placed away from the centre of the field of view [24, 30]. All these factors will also affect image quantification and it is important to mention that the calibration factor obtained when placing an animal in the centre of the field-of-view might not be valid for axially displaced animals. Therefore,

it is crucial for labs that frequently use multiple-animal imaging to further evaluate this aspect in order to obtain accurate and reliable quantitative data.

We hope with these guidelines that many preclinical imaging labs will implement a QC programme for their PET and SPECT scanner. But we are also convinced that it is worth going one step further and working towards an accreditation programme for PET and SPECT, similar to the EARL initiative for clinical PET scanners (http://earl.eanm.org), which might open the way towards harmonization initiatives in preclinical imaging.

Liability Statement

This guideline summarizes the views of the EANM Physics Committee, the EANM Translational Molecular Imaging & Therapy Committee, and the ESMI STANDARD group. It reflects recommendations for which the EANM/ESMI cannot be held responsible. The recommendations should be taken into context of good preclinical practice and do not substitute for national and international legal or regulatory provisions.

Author Contributions All authors contributed to the study conception and design. The first draft of the manuscript was written by Christian Vanhove and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript. This manuscript was also reviewed by the EANM Committees, the EANM National Delegates and the ESMI Executive Committee.

Data Availability The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval This is guideline for implementing an efficient quality control programme for preclinical PET and SPECT that requires no ethical approval.

Consent to participate This is guideline for implementing an efficient quality control programme for preclinical PET and SPECT that requires no written informed consent.

Consent to publish Not applicable

Competing interests All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this guideline.

References

1. Lauber DT, et al. State of the art in vivo imaging techniques for laboratory animals. Lab Anim. 2017;51:465–78.
2. Kiessling F, Pichler BJ. Small animal imaging: Basics and practical guide. Small Anim Imaging Basics Pract Guide. 2011. https://doi.org/10.1007/978-3-642-12945-2.
3. Summary Report on the statistics on the use of animals for scientific purposes in the Member States of the European Union and Norway in 2018. 2018. https://ec.europa.eu/environment/chemicals/lab_animals/reports_en.htm.
4. James ML, Gambhir SS. A molecular imaging primer: modalities, imaging agents, and applications. Physiol Rev. 2012;92:897–965.
5. Cunha L, et al. Preclinical imaging: an essential ally in modern biosciences. Mol Diagnosis Ther. 2014;18:153–73.
6. Vanhove C, et al. Accurate molecular imaging of small animals taking into account animal models, handling, anaesthesia, quality control and imaging system performance. EJNMMI Phys. 2015;2:31.
7. Russell WM, Burch R. Principle of Human Experimental Techniques. 1959.
8. Tannenbaum J, Bennett BT. Russell and Burch's 3Rs then and now: The need for clarity in definition and purpose. J Am Assoc Lab Anim Sci. 2015;54:120–32.
9. Van Der Heyden B, Van Hoof SJ, Schyns LE, Verhaegen F. The influence of respiratory motion on dose delivery in a mouse lung tumour irradiation using the 4D MOBY phantom. Br J Radiol. 2017;90. https://doi.org/10.1259/bjr.20160419.
10. Caruana CJ, et al. EFOMP policy statement 16: the role and competences of medical physicists and medical physics experts under 2013/59/EURATOM. Phys Medica. 2018;48:162–8.
11. McDougald W, et al. Standardization of preclinical PET/CT imaging to improve quantitative accuracy, precision, and reproducibility: a multicenter study. J Nucl Med. 2020;61:461–8.
12. Dillenseger JP, Choquet P, Snay E, Fragoso Costa P. Why the preclinical imaging field needs nuclear medicine technologists and radiographers? Eur J Hybrid Imaging. 2020;4. https://doi.org/10.1186/s41824-020-00081-z.
13. Fragoso Costa P, Santos A, Testanera G. An insight into the EANM technologist committee benchmark document on nuclear medicine technologists' competencies. Eur J Nucl Med Mol Imaging. 2017;44:1604–6.
14. Tavares AAS, Mezzanotte L, Mcdougald W, Bernsen MR, Vanhove C. Community survey results show that standardisation of preclinical imaging techniques remains a challenge. Mol Imaging Biol. 2022. https://doi.org/10.1007/s11307-022-01790-6.
15. McDougald WA, Mannheim JG. Understanding the importance of quality control and quality assurance in preclinical PET/CT imaging. EJNMMI Phys. 2022;9. https://doi.org/10.1186/s40658-022-00503-w.

16. Mannheim JG, et al. Standardization of small animal imaging—current status and future prospects. Mol Imaging Biol. 2018;20:716–31.

17. Association NEMA. NEMA Standard Publication NU 4–2008: performance measurements of small animal positron emission tomographs. Rosslyn: National Electrical Manufacturers Association; 2008.

18. Goertzen AL, et al. NEMA NU 4–2008 comparison of preclinical PET imaging systems. J Nucl Med. 2012;53:1300–9.

19. Bao Q, Newport D, Chen M, Stout DB, Chatziioannou AF. Perfrmance evalution of the inveon dedicated PET preclinical tomograph based on the NEMA NU-4 standards. J Nucl Med. 2009;50:401–8.

20. Krishnamoorthy S, Blankemeyer E, Mollet P, Surti S, Van Holen R, Karp JS. Performance evaluation of the MOLECUBES β-CUBE - A high spatial resolution and high sensitivity small animal PET scanner utilizing monolithic LYSO scintillation detectors. Phys Med Biol. 2018;63. https://doi.org/10.1088/1361-6560/aacec3.

21. Disselhorst JA, et al. Image-quality assessment for several positron emitters using the NEMA NU 4–2008 standards in the siemens inveon small-animal PET scanner. J Nucl Med. 2010;51:610–7.

22. Halle P, Schug D, Schulz V. Comments on the NEMA NU 4–2008 standard on performance measurement of small animal positron emission tomographs. EJNMMI Phys. 2020;7. https://doi.org/10.1186/s40658-020-0279-2.

23. Osborne DR, Kuntner C, Berr S, Stout D. Guidance for efficient small animal imaging quality control. Mol Imaging Biol. 2017;19:485–98.

24. Prieto E, et al. Performance evaluation of a preclinical SPECT/CT system for multi-animal and multi-isotope quantitative experiments. Sci Rep. 2022;12:1–13.

25. Grigg OA, Farewell VT, Spiegelhalter DJ. Use of risk-adjusted CUSUM and RSPRT charts for monitoring in medical contexts. Stat Methods Med Res. 2003;12:147–70.

26. Park MA, et al. Adsorption of metallic radionuclides on plastic phantom walls. Med Phys. 2008;35:1606–10.

27. BusemannSokole E, et al. Routine quality control recommendations for nuclear medicine instrumentation. Eur J Nucl Med Mol Imaging. 2010;37:662–71.

28. Rousset OG, Ma Y, Evans AC. Correction for partial volume effects in pet: principle and validation. J Nucl Med. 1998;39:904–11.

29. Prieto E, et al. Evaluation of spatial resolution of a PET scanner through the simulation and experimental measurement of the recovery coefficient. Comput Biol Med. 2010;40:75–80.

30. Greenwood HE, Nyitrai Z, Mocsai G, Hobor S, Witney TH. High-throughput PET/CT imaging using a multiple-mouse imaging system. J Nucl Med. 2020;61:292–7.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Christian Vanhove[1] · Michel Koole[2] · Pedro Fragoso Costa[3] · Margret Schottelius[4] · Julia Mannheim[5] · Claudia Kuntner[6] · Geoff Warnock[7,8] · Wendy McDougald[9,10] · Adriana Tavares[9,11] · Monique Bernsen[12]

✉ Christian Vanhove
Christian.Vanhove@UGent.be

Michel Koole
michel.koole@kuleuven.be

Pedro Fragoso Costa
Pedro.FragosoCosta@uk-essen.de

Margret Schottelius
Margret.Schottelius@chuv.ch

Julia Mannheim
julia.mannheim@med.uni-tuebingen.de

Claudia Kuntner
claudia.kuntner-hannes@meduniwien.ac.at

Geoff Warnock
Geoffrey.Warnock@bruker.com

Wendy McDougald
wendy.mcdougald@siemens-healthineers.com

Adriana Tavares
adriana.tavares@ed.ac.uk

Monique Bernsen
m.bernsen@erasmusmc.nl

[1] Department of Electronics and Information Systems, Faculty of Engineering and Architecture, Campus UZ Gent, Institute Biomedical Engineering and Technology (IBiTech), Corneel Heymanslaan 10, 9000 Ghent, Belgium

[2] Nuclear Medicine and Molecular Imaging, KU Leuven, Louvain, Belgium

[3] Department of Nuclear Medicine, University Hospital Essen, West German Cancer Center (WTZ), University of Duisburg-Essen, Essen, Germany

[4] Unit of Translational Radiopharmaceutical Sciences, Departments of Nuclear Medicine and of Oncology, Centre Hospitalier Universitaire Vaudois (CHUV) and University of Lausanne (UNIL), Lausanne, Switzerland

[5] Department of Preclinical Imaging and Radiopharmacy, Werner Siemens Imaging Center, Eberhard-Karls University Tübingen, Tübingen, Germany

[6] Department of Biomedical Imaging and Image-Guided Therapy, Medical University of Vienna, Vienna, Austria

[7] University of Zurich, Zurich, Switzerland

[8] PMOD Technologies LLC, Fällanden, Switzerland

[9] BHF-University Centre for Cardiovascular Science, University of Edinburgh, Edinburgh, UK

[10] Siemens, Molecular Imaging, Hoffman Estates,, IL, USA

[11] Edinburgh Imaging, University of Edinburgh, Edinburgh, UK

[12] AMIE Core Facility, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands