Multicenter survey of PET/CT protocol parameters that affect standardized uptake values

Darrin Byrd
Rebecca Christopfel
John Buatti
Eduardo Moros
Sadek Nehmeh
Adam Opanowski
Paul Kinahan

Multicenter survey of PET/CT protocol parameters that affect standardized uptake values

Darrin Byrd,^a Rebecca Christopf,^a John Buatti,^b Eduardo Moros,^c Sadek Nehmeh,^d Adam Opanowski,^e and Paul Kinahan^{a,*}

^aUniversity of Washington, Department of Radiology, Seattle, Washington, United States

^bUniversity of Iowa, Department of Radiation Oncology, Iowa City, Iowa, United States

^cMoffitt Cancer Center, Tampa, Florida, United States

^dCornell University, Weill Cornell Medical College, New York, New York, United States

^eAmerican College of Radiology, Philadelphia, Pennsylvania, United States

Abstract. Clinical trials that evaluate cancer treatments may benefit from positron emission tomography (PET) imaging, which for many cancers can discriminate between effective and ineffective treatments. However, the image metrics used to quantify disease and evaluate treatment may be biased by many factors related to clinical protocols and PET system settings, many of which are site- and/or manufacturer-specific. An observational study was conducted using two surveys that were designed to record key sources of bias and variability in PET imaging. These were distributed to hospitals across the United States. The first round of surveys was designed and distributed by the American College of Radiology's Centers of Quantitative Imaging Excellence program in 2011. The second survey expanded on the first and was completed by the National Cancer Institute's Quantitative Imaging Network. Sixty-three sites responded to the first survey and 36 to the second. Key imaging parameters varied across participating sites. The range of reported methods for image acquisition and reconstruction suggests that signal biases are not matched between sites. Patient preparation was also inconsistent, potentially contributing additional variability. For multicenter clinical trials, efforts to control biases through standardization of imaging procedures should precede patient measurements. © 2017 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: [10.1117/1.JMI.5.1.011012](https://doi.org/10.1117/1.JMI.5.1.011012)]

Keywords: PET quantitation; PET clinical trial; standardized uptake value; multicenter PET.

Paper 17203SSR received Jul. 7, 2017; accepted for publication Nov. 1, 2017; published online Dec. 8, 2017.

1 Introduction

Positron emission tomography combined with x-ray computed tomography (PET/CT) imaging has become a standard component of oncology diagnosis and staging as well as a tool to plan and assess response to therapy. PET image metrics such as the standardized uptake value (SUV) have been shown to be effective at characterizing disease *in vivo*. This information may potentially spare patients the side effects of ineffective treatment,¹ reduce unneeded invasive surgeries,^{2–4} and allow for more efficient clinical trials aimed at discovering new therapies. Quantitative PET/CT using the glucose analogue 18F-fluorodeoxyglucose (FDG) is a valuable tool for assessment of an individual's response to therapy because SUVs can measure metabolic changes, which can be a more useful indicator of response than anatomical size changes.⁵ However, the use of PET in this task is complicated by signal bias and variability that result from inherent characteristics of PET as well as choices of parameters made by clinicians and scanner operators.

The SUV is a relative measure of FDG uptake that roughly compensates for the two most significant factors that affect tracer concentration in patients: the injected radiotracer dose and patient size.⁶ The test–retest variability of SUVs for FDG-PET oncology imaging has been shown to be 10% to 12% in carefully controlled single-center studies,^{7,8} though it may be larger

in more typical clinical conditions.⁹ In studies conducted across multiple sites, variability is greater, having been measured at 18% in some studies¹⁰ and 40% to 50% in others.^{9,11}

This increased variability between hospitals is likely due to intersite variability of factors that influence SUV biases. As Fig. 1 shows, the sources of error include the independent effects of the imaging physics and underlying patient variability, inconsistent patient preparation and scan protocols, and data processing and image reconstruction.^{12–14} Multicenter trials have the advantage of faster patient accrual versus single-center trials but increases in SUV variability can reduce study power even as the number of patients studied increases.¹⁵

Surveys of PET imaging protocols that aimed to characterize this intersite variability of PET practices were reported by Graham et al.¹⁶ and Beyer et al.¹⁷ Independently, each group prepared and submitted a survey to centers using FDG-PET imaging in clinical practice.

This report shares the results of two subsequent surveys. In these cases, however, the surveys were specifically aimed at quantitative imaging protocols used for oncology clinical trials. The first survey was conducted by the American College of Radiology Imaging Network (ACRIN) in 2011 as part of the initiation of the Centers of Quantitative Imaging Excellence (CQIE) qualification program.¹⁸ The primary objective of the CQIE program was to establish a network of “trial ready”

*Address all correspondence to: Paul Kinahan, E-mail: kinahan@uw.edu

Fig. 1 Data flow and contributions of variance to overall SUV variability. Note that calibrations take place in parallel with clinical measurements, but at their own site-specific frequency.

sites within the National Cancer Institutes' Cancer Centers Program capable of conducting clinical trials with integral molecular and/or functional advanced imaging endpoints. The second survey was conducted in 2013 by the Quantitative Imaging Network (QIN), a network of grant-supported sites using advanced cancer imaging in clinical trials.¹⁹

2 Methods

2.1 Centers of Quantitative Imaging Excellence Survey

The CQIE survey was initiated in 2011 as a component of the 3-year CQIE qualification program. Participating cancer centers underwent an initial qualification assessment that included a survey, phantom scans, assessment of clinical images, and a standardized set of quality control procedures. The CQIE qualification program included PET, CT, and MRI scanners at National Cancer Institute Designated Cancer Centers. Sites were required to submit the data reported here in order to be accredited by the CQIE.²⁰ The CQIE survey (Fig. 2) consisted of eight questions, some with subquestions, and was repeated for both body- and brain-imaging protocols used in clinical trials at each site. Sites were allowed to submit data on multiple PET scanners, and information on PET scanner manufacturers and models was collected separately.

1. Reconstruction algorithm:	
2. Field-of-view:	cm
3. Slice thickness:	mm
4. Pixel size:	mm
5. Matrix size:	x
6. Data corrections:	
a. Attenuation:	
☐	CT
☐	CT-segment
b. Randoms:	
☐	Subtract (online)
☐	Smoothed (offline)
c. Scatter algorithm:	
7. For OSEM:	
a. Iterations:	
b. Subsets:	
8. Post-processing:	
a. Filter type:	
b. Setting:	
☐	None

Fig. 2 CQIE PET imaging survey.

ACQUISITION	Brain protocol	Body protocol
Min. prescan fasting (hrs)		
Low carb diet? / compliance		
Max. blood Glu level		
Amount FDG injected		
Post-inj. assay? (Y/N)		
Uptake time (target)		
Typical uptake time range		
Mode (2-D or 3-D)		
Time per bed position		
bed overlap (cm or slices)		
Other acquisition info		
PROCESSING		
FOV diameter (cm)		
Slice Thickness (mm)		
Pixel size (mm)		
Matrix size (N x N)		
Attenuation (CT, kVP, seg.?)		
Randoms (smoothed, online?)		
Scatter (model, subtract, in loop?)		
Recon. Algorithm		
Iterative recon settings		
FBP recon settings		
Other processing info		

Fig. 3 QIN PET imaging survey.

2.2 Quantitative Imaging Network Survey

The QIN PET survey was initiated in 2013 and was designed based on the results of the CQIE survey. The survey (Fig. 3) consisted of 22 questions that were repeated for both body- and brain-imaging protocols used in clinical trials at each site. Sites were encouraged to submit data on multiple PET scanners. PET scanner manufacturers and years of installation were also recorded.

Both surveys asked about clinical trial protocols and did not record differences, if any existed, between clinical trial imaging and routine protocols.

3 Results

The CQIE survey was sent to 55 hospitals and a total of 63 unique PET/CT scanner responses were analyzed.

Due to some sites' involvement in both CQIE and QIN groups, some sites submitted multiple instances of the survey. In most cases, the responses were equivalent and duplicate answers were discarded. However, if answers in redundant submissions differed but were both apparently valid (for instance, a 55- and 70-cm field of view), both were included in the analysis. Not all sites answered all questions.

The QIN survey was distributed to the QIN's Data Acquisition Working Group as well as those sites that had participated in the CQIE survey 2 years earlier. Thirty-six hospitals responded, and a total of 44 unique PET/CT scanner responses were received. Most sites had a single PET/CT scanner; eight sites had two scanners. Approximately 75% of the sites responding to the QIN survey had been included in the CQIE survey years earlier.

3.1 Scanner Type

Of the CQIE-survey scanners, 33 were General Electric, 23 were Siemens, and 7 were Philips. Of the 44 scanners included in the QIN survey, 24 were General Electric, 18 were Siemens, and 2 were Philips. The average year of installation for QIN-survey

Fig. 4 Histogram of scanner age for QIN-survey sites in 2013.

Table 1 Image matrix sizes. All sites used square matrices for image reconstruction.

Size ($n \times n$)	Body imaging		Brain imaging	
	CQIE	QIN	CQIE	QIN
128	34	26	36	19
144	7	3	1	1
168	14	9	1	2
172		1	13	17
192	3	1	8	3
200	3	4		1
256	1	4		1
400		1	1	4
512		1		2
No answer	1		3	
Total	63	50	63	50

scanners was 2008, with a standard deviation of 2.9 years. This corresponds to an average age of ~ 6 years at the time the survey was conducted (Fig. 4).

3.2 Reconstruction and Processing

For scanners in the QIN survey, the reported trans-axial field-of-view diameter was 63 ± 11 cm (range 30 to 81 cm) for body imaging and 32 ± 10 cm (range 25 to 70 cm) for brain imaging. Table 1 shows the reported matrix sizes for both surveys.

For body imaging, the average trans-axial voxel size was 4.2 ± 1.2 mm. For brain, it was 2.0 ± 1.1 mm. Slice thicknesses had distributions of 3.57 ± 0.89 mm and 2.98 ± 0.71 mm for body and brain, respectively. Nineteen scanners used the same slice thickness for brain and body. The distributions of trans-axial voxel dimensions and slice thicknesses are shown in Fig. 5.

Reconstruction methods varied across scanners in the QIN survey. The most common algorithm reported was the ordered-subsets expectation maximization (OSEM) algorithm²¹ without time-of-flight data, used by 31 sites (one of these used two-dimensional OSEM). Eight sites indicated their reconstructions used time-of-flight data and three used the analytic three-dimensional-reprojection method.²² Some sites used Fourier rebinning.

Figures 6 and 7 show the reported image smoothing parameters versus the number of iterative updates (defined as the number of iterations times the number of subsets). On the x -axis, where a postfilter of 0 mm is indicated, no postfiltering was applied. Algorithms plotted on the y -axis, implying no updates, are noniterative methods.

3.3 Injected Radioactivity

Of the QIN-survey sites, 24 provided a fixed value of injected radiotracer for body imaging of adults, 11 provided a range, and 15 reported a weight-based injection. Fixed injections and injection ranges in mCi are shown in Tables 2 and 3. For sites reporting weight-based injections for body scans, injections were computed with coefficients between 0.14 and 0.20 mCi/kg (average of 0.15 mCi/kg). Three sites did not report their method of calculating weight-based injections for body scans and one site reported using (BMI)/3 mCi, where BMI is the body mass index.

Fig. 5 Histograms of (a) trans-axial voxel sizes and (b) slice thicknesses for body and brain imaging.

Fig. 6 For body imaging, smoothing (postfilter full width at half maximum) versus iterative updates for the (a) QIN survey and (b) CQIE survey. Where multiple points lie in the same location, small shifts have been introduced to make markers visible.

Fig. 7 For brain imaging, smoothing (postfilter full width at half maximum) versus iterative updates for the (a) QIN survey and (b) CQIE survey. Where multiple points lie in the same location, small shifts have been introduced to make markers visible.

Table 2 FDG injections for body imaging for sites reporting a fixed value.

Injection (mCi)	<i>n</i>
7	2
8	1
10	4
12	6
13	1
14	2
15	8
Total	24

Table 3 FDG injections for body imaging for sites reporting a range of values

Injection (mCi)	<i>n</i>
5 to 10	1
5 to 15	2
6 to 15	1
7 to 12	3
10 to 15	3
10 to 20	1
Total	11

Table 4 FDG injections for brain imaging for sites reporting a fixed value.

Injection (mCi)	<i>n</i>
5	6
7	2
8	4
10	13
11	1
12	4
13	1
15	4
Total	35

The injection for brain scans was on average less than that for body scans. Three sites responded with a range of administered tracer injections for brain scans: two reported 5 to 10 mCi and one reported 6 to 10 mCi. Thirty-five sites reported a fixed injection for brain scans (Table 4). The remaining sites used a weight-based injection, with coefficients in mCi/kg of 0.10, 0.14, 0.18, and 0.20. Each value was reported by one site except for 0.14 mCi/kg, which was reported by two.

3.4 Prescan Fasting

All QIN-survey sites required prescan fasting for both body and brain imaging. The required periods are shown in Tables 5 and 6.

For brain imaging, the majority of QIN-survey sites followed the same fasting period as body imaging with 17 sites requiring a 4-h fasting period and 18 sites requiring 6 h. Three sites required one half-hour of fasting, one required 12 h, and 2 sites required 4 to 6 h. Four sites differentiated diabetic and nondiabetic patients for brain scans, all requiring 6 and 4 h, respectively, as for body scans.

For body and brain scans, carbohydrate restriction was the same at each QIN-survey site. That is, responses were identical between body and brain protocols. Nineteen sites required

Table 5 Required period of prescan fasting for body imaging.

Duration (h)	<i>n</i>
1	3
4	16
6	19
4 (6 for diabetics)	5
Other	2
Total	45

Table 6 Required period of prescan fasting for brain imaging.

Duration (h)	<i>n</i>
0.5	3
4	17
6	18
4 (6 for diabetics)	4
12	1
4 to 6	2
Total	45

low-carbohydrate compliance, whereas 24 did not. The remaining two sites required low-carbohydrate compliance only for diabetics.

3.5 Blood Glucose Levels

Twenty-nine QIN-survey sites reported 199 or 200 mg/dL as the maximum blood glucose level allowable in order to proceed with injection and scanning for body imaging. All reported limits are shown in Table 7.

For brain imaging, cutoffs were similar to body imaging. All values are shown in Table 8.

Table 7 Maximum allowable blood glucose for body imaging.

Concentration (mg/dL)	<i>n</i>
150	4
199	2
200	27
250	3
Other	9
Total	45

Table 8 Maximum allowable blood glucose for brain imaging.

Concentration (mg/dL)	<i>n</i>
150	4
199	2
200	25
250	5
Other	9
Total	45

Table 9 QIN-survey sites' target uptake times.

Uptake time (min)	N (body)	N (brain)
30		12
40		1
45	3	1
50		3
60	31	19
75	1	
90	7	
Other	3	6

Table 10 QIN-survey sites' reported typical ranges of actual uptake times relative to target times.

Variation in uptake time (min)	Body	Brain
Target = actual	24	27
±5	1	1
±10	6	4
+5	3	
+10	5	5
−15	2	1
+15	6	8
+15 to 30	2	
+ > 30	1	3

3.6 Uptake Time

Uptake time was quantified by two measures. Sites reported their clinical protocols' prescribed time intervals between patient injection and scan start as their "target" uptake time, as well as a "typical range" that reflected the variability introduced by busy clinical schedules and outside constraints.

Target uptake times are shown in Table 9. Table 10 shows sites' self-reported compliance. For body imaging, the most common uptake time was 60 min, and nearly half of the sites reported no difference between targeted and actual uptake times.

For brain imaging, the most common time was 30 min and again approximately half of the sites reported that targeted and actual uptake times were equal.

3.7 Bed Position Duration

Some QIN-survey sites provided discrete values for bed position timing and others provided a range. Durations are shown in Table 11. Fixed durations ranged from 1 to 10 min, and the distribution was 3.5 ± 1.6 min. One site (but two different scanners) provided a range of 2 to 5 min based on BMI.

Table 11 Scan duration per bed position for body imaging at QIN-survey sites.

Duration (min)	N
1 to 3	1
1.5 to 3	1
1.5 to 4	2
2 to 3	4
2 to 5	4
3 to 4	1
3 to 5	1
3 to 10	1
1	1
1.5	1
2	2
2.5	2
3	15
4	6
4.25	1
6	1
7	1
10	1
Total	46

For brain scans, only two sites reported a range. One used 8 to 10 min and the other used a BMI-based duration. The distribution of fixed durations was 11.7 ± 5.9 min (range of 4 to 30 min). Two sites reported that they used multiple bed positions. Durations are shown in Table 12.

Correlations were not observed between injected dose and scan duration for body or brain imaging. For body dose versus body duration and for brain dose versus brain duration, the correlation coefficient was the same: -0.25 .

4 Discussion

An array of previous studies has reported on the effects on SUV bias and variability in PET imaging procedures. The CQIE, and to a greater extent, the QIN, created their surveys in response to these data in order to characterize the range of key parameters used for clinical trial imaging in a large group of hospitals. In principle, the variability observed in these imaging parameters could be translated into variability of PET image metrics. However, we note that the parameter space is large, and therefore, such analysis risks drawing conclusions that may be valid only for a subset of parameters or, with sufficient parameter mismatch, not valid at all. With appropriate caution, the discussion below attempts to place the variations seen in our survey data in context using previous PET signal studies that employed

Table 12 Scan duration per bed position for brain imaging at QIN-survey sites.

Minutes	<i>n</i>
8 to 10	1
4	1
5	3
6	2
8	4
10	16
12	1
15	7
20	4
30	2
10 or 20	1
BMI-based	1
Total	44

intentional changes in the imaging process. We focus in this study on SUVs, but we note that metabolic tumor volume, another proposed biomarker, will also be affected by the variations of imaging parameters described below.

PET scanners from three manufacturers were used by sites in this survey. This allows for a wide variation in detection and reconstruction methods, spanning regular and time-of-flight reconstructions, disparate ways of modeling the point-spread-function (if it is modeled at all), and a spectrum of user-specified smoothness settings (Figs. 6 and 7). These factors can influence reconstructed image resolution, and with it, the biases that affect small lesions.^{12,23–26} Changes in SUV as large as 26% have been seen when final image resolution changes by <4 mm.²⁵ For reference, filtering an image that has 5-mm resolution with a 7.5-mm Gaussian kernel will produce a 4-mm change in resolution (using addition in quadrature). We note these numbers, 5 and 7.5 mm, roughly correspond to typical PET scanner resolution without filtering²⁷ and the range of filter widths shown in Figs. 6 and 7.

We note that it is possible that differing numbers of image updates could mitigate some of the effects of differing postfilters described above, as iterations have also been shown to have an effect on SUV bias.²⁸ That is, simultaneously reducing the number of updates and the postfilter width could in principle leave bias fixed. However, the lack of correlations between updates and filtering in Figs. 6 and 7 suggests that these parameters are not being selected in such a manner.

Even with matched detection and postreconstruction image filtering, changes in trans-axial voxel size can contribute to changes in signal bias similar in magnitude to the resolution effects described above.²⁹ The reported range of voxel sizes, which spans more than a factor of 2 (Fig. 5), is another potentially significant source of variability in SUVs across survey sites.

Uptake time was included as a key imaging parameter because SUVs for some cancers have been shown to vary approximately linearly over a finite interval of time after injection. In particular, for the period of 27 to 75 min postinjection, Beaulieu et al.³⁰ saw SUVs vary by up to 0.15 and 0.12 min^{−1} for maximum and mean SUVs, respectively, (maximum SUV is computed using the largest single voxel value in a region while mean SUV is computed using a region's average value). The reported anecdotal ranges of ± 15 min with respect to the target times estimated by sites would correspond to changes of up to ± 2.3 and ± 1.8 for maximum and mean SUVs. Furthermore, the overall range of target times (40 to 90 min) means that large SUV bias differences could exist between sites even if target times are met consistently.

Variability in injected dose and scan duration both affect image noise, because they lead to variations in the total number of detected events. It is known that decreasing detected events increases image noise. This can affect the bias and variability of maximum SUVs. It can also lead to increased variability of mean SUVs.^{12,29} We found that scan durations differed by a factor of 4 and did not find that injected dose was correlated with scan duration.

Blood glucose can significantly suppress SUVs if it is uncontrolled. Large changes in SUVs have been reported even in patient populations that would be in compliance with the limit of 200 mg/dL allowed at many sites.³¹ The true impact of blood glucose is difficult to estimate from this survey, as the actual distribution of blood glucose in patients is not known. However, fasting guidelines were noted to vary from 0.5 to 12 h, suggesting standardization is needed in patient protocols.

Although the surveys covered many key factors in SUV bias, they were not exhaustive. The surveys did not ask about patient- or disease-specific protocols, respiratory gating methods, or SUV computation. The surveys did not mention that the software used or analysis done by physicians, areas that also could benefit from standardization.³² Although reconstruction parameters were reported, their meanings are not standardized across scanner types. This is particularly true of those that describe OSEM: detection modeling, attenuation, and scatter correction, as well as postfiltering may all be done in proprietary ways. Consequently, surveys are limited in their ability to estimate the effects of image resolution on SUV variability.

Another limitation of this study is in how the surveys were conducted. The first CQIE data collection was not designed to capture all the parameters that affect SUVs. However, we feel this is important data to include. The second QIN survey was designed to capture all parameters that affect SUVs. No surveying methodology was used to design the surveys as we collected objective parameters, i.e., no opinions or statistical samplings were collected. The overlap of participating sites between the two surveys means the data sets are not independent, but this is not a confounding factor in relation to the qualitative conclusion that standardization of imaging protocols is needed.

The surveys did not ask about scanner calibration methods or frequency, which play a significant role in SUV accuracy.³³ However, as a follow up to this survey, the QIN Data Acquisition Working Group implemented a multicenter calibration test that involved regular measurements of scanner and dose calibrator biases at QIN sites. These data will help to characterize the stability these instruments' calibration across a national network of research hospitals.

Many studies have been completed or are underway regarding the repeatability of PET metrics. They employ varying methods, such as surveys, phantom measurements, and patient scans. We note that phantom scans are an integral part of site qualification for clinical trials^{34,35} and accreditation, as they provide key information that cannot be obtained from surveys. The variability seen in this work demonstrates that imaging protocols must also be evaluated. To that end, we believe published guidelines such as the Uniform Protocol for Imaging in Clinical Trials³⁶ or the European Association of Nuclear Medicine guidelines³⁷ also have clear value. Additionally, we believe that the utilization of new phantom techniques, which aim to estimate image resolution from phantom scans,^{38–40} may enable the selection of imaging parameters that lead to better-matched resolution effects, which are still a significant hurdle in achieving uniform bias for small lesions.

Where our surveys overlap with past literature, there is general agreement that standardization of protocol parameters is lacking.^{16,17} However, a rigorous characterization of how factors affecting SUV bias might vary across hospitals generally (and not within a trial-focused subset) has never been offered in the literature, likely because of the considerable number factors to be recorded and the prevailing view that qualitative imaging is presently sufficient for most hospitals using PET/CT.

While there is no universally accepted threshold for differences in SUVs in clinical trials, the PERCIST criteria propose a difference of $\pm 30\%$ as indicating a true change.⁴¹ This can be compared to the QIBA FDG-PET/CT profile,⁴² which states that “[a] measured increase in [maximum SUV] of 39% or more, or a decrease of -28% or more, indicates that a true change has occurred with 95% confidence.” We note that the specifications for an individual clinical trial, or a clinical indication, may be set using other criteria, but the general guidelines suggest that the imaging parameter variations described above are plausibly large enough to impact evaluation of cancer therapies.

5 Conclusion

We have presented the findings of two related surveys on clinical variations in PET/CT acquisition and processing. Based on the range of reported imaging parameters, we expect that large differences in biases may exist among the sites that participated in these surveys. This was an observational study, and we cannot directly estimate the combined impact of the parameter variations on the net variations in PET SUVs. We feel that recommending optimal parameters is beyond the scope of this work. However, one reasonable interpretation of our results is that standardization of the parameters above should precede any multicenter trial that uses PET SUVs quantitatively. This should be a high priority for future multicenter trials using quantitative imaging.

Disclosures

No relevant financial interests were disclosed.

Acknowledgments

We would like to thank all members of the QIN-DAWG team and ACRIN members who participated in this survey. The participants in the survey include from the QIN: Massachusetts General Hospital Martinos Center, Memorial Sloan-Kettering Cancer Center, University of Pittsburgh, Moffit Cancer Center, University of Iowa, Ohio State University, University of

Washington, and Johns Hopkins University, University of Pennsylvania and from CQIE: UC San Francisco, University of Chicago, Memorial Sloan-Kettering Cancer Center, City of Hope, Dana Farber Cancer Institute, UT Southwestern, University of Utah, Cancer Institute of New Jersey, Fox Chase Cancer Center, Johns Hopkins University, Children’s Hospital of Michigan PET Center, MD Anderson Cancer Center, Moffit Cancer Center, UC Davis, UC Irvine, University of Michigan, University of Pittsburgh, VCU Massey, and Medical University of South Carolina. Grant Nos. NCI 1 U01 CA143062, U01-CA148131, U10CA180820, and NIH Contract No. HHSN261200800001E

References

1. W. A. Weber, “Use of PET for monitoring cancer therapy and for predicting outcome,” *J. Nucl. Med.* **46**, 983–995 (2005).
2. B. M. Fischer et al., “Multimodality approach to mediastinal staging in non-small cell lung cancer. Faults and benefits of PET-CT: a randomised trial,” *Thorax* **66**, 294–300 (2011).
3. H. van Tinteren et al., “Effectiveness of positron emission tomography in the preoperative assessment of patients with suspected non-small-cell lung cancer: the PLUS multicentre randomised trial,” *Lancet* **359**, 1388–1392 (2002).
4. D. E. Maziak et al., “Positron emission tomography in staging early lung cancer: a randomized trial,” *Ann. Intern. Med.* **151**, 221–228 (2009).
5. W. A. Weber, “Assessing tumor response to therapy,” *J. Nucl. Med.* **50**(Suppl. 1), 1S–10S (2009).
6. J. A. Thie, “Understanding the standardized uptake value, its methods, and implications for usage,” *J. Nucl. Med.* **45**, 1431–1434 (2004).
7. H. Minn et al., “Lung cancer: reproducibility of quantitative measurements for evaluating 2-[F-18]-fluoro-2-deoxy-D-glucose uptake at PET,” *Radiology* **196**, 167–173 (1995).
8. W. A. Weber et al., “Reproducibility of metabolic measurements in malignant tumors using FDG PET,” *J. Nucl. Med.* **40**(11), 1771–1777 (1999).
9. V. Kumar et al., “Variance of standardized uptake values for FDG-PET/CT greater in clinical practice than under ideal study settings,” *Clin. Nucl. Med.* **38**, 175–182 (2013).
10. L. M. Velasquez et al., “Repeatability of 18F-FDG PET in a multicenter phase I study of patients with advanced gastrointestinal malignancies,” *J. Nucl. Med.* **50**, 1646–1654 (2009).
11. F. H. Fahey et al., “Variability in PET quantitation within a multicenter consortium,” *Med. Phys.* **37**, 3660–3666 (2010).
12. R. Boellaard, “Standards for PET image acquisition and quantitative data analysis,” *J. Nucl. Med.* **50**, 11S–20S–20S (2009).
13. M. C. Adams et al., “A systematic review of the factors affecting accuracy of SUV measurements,” *Am. J. Roentgenol.* **195**, 310–320 (2010).
14. P. E. Kinahan and J. W. Fletcher, “PET/CT standardized uptake values (SUVs) in clinical practice and assessing response to therapy,” *Semin. Ultrasound, CT MRI* **31**, 496–505 (2010).
15. R. K. Doot et al., “Design considerations for using PET as a response measure in single site and multicenter clinical trials,” *Acad. Radiol.* **19**, 184–190 (2012).
16. M. M. Graham, R. D. Badawi, and R. L. Wahl, “Variations in PET/CT methodology for oncologic imaging at US academic medical centers: an imaging response assessment team survey,” *J. Nucl. Med.* **52**, 311–317 (2011).
17. T. Beyer, J. Czernin, and L. S. Freudenberg, “Variations in clinical PET/CT operations: results of an international survey of active PET/CT users,” *J. Nucl. Med.* **52**, 303–310 (2011).
18. M. Rosen et al., “Performance observations of scanner qualification of NCI-designated cancer centers: results from the Centers of Quantitative Imaging Excellence (CQIE) program,” *Acad. Radiol.* **24**, 232–245 (2017).
19. L. P. Clarke et al., “The Quantitative Imaging Network: NCI’s historical perspective and planned goals,” *Transl. Oncol.* **7**, 1–4 (2014).
20. J. S. Scheuermann et al., “Qualification of national cancer institute-designated cancer centers for quantitative PET/CT imaging in clinical trials,” *J. Nucl. Med.* **58**, 1065–1071 (2017).

21. H. M. Hudson and R. S. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE Trans. Med. Imaging* **13**(4), 601–609 (1994).
22. P. E. Kinahan and J. G. Rogers, "Analytic 3D image reconstruction using all detected events," *IEEE Trans. Nucl. Sci.* **36**, 964–968 (1989).
23. D. Byrd et al., "A post-hoc methodology for harmonizing PET image reconstruction protocols," *J. Nucl. Med.* **55**, 372 (2014).
24. J. Sunderland et al., "Assessment of PET/CT reconstruction harmonization through Gaussian post-filtration," *J. Nucl. Med.* **57**, 1917 (2016).
25. M. Westertorp et al., "Quantification of FDG PET studies using standardised uptake values in multi-centre trials: effects of image reconstruction, resolution and ROI definition parameters," *Eur. J. Nucl. Med. Mol. Imaging* **34**, 392–404 (2007).
26. N. C. Krak et al., "Effects of ROI definition and reconstruction method on quantitative outcome and applicability in a response monitoring trial," *Eur. J. Nucl. Med. Mol. Imaging* **32**, 294–301 (2005).
27. M. E. Daube-Witherspoon et al., "PET performance measurements using the NEMA NU 2-2001 standard," *J. Nucl. Med.* **43**, 1398–1409 (2002).
28. C. J. Jaskowiak et al., "Influence of reconstruction iterations on 18F-FDG PET/CT standardized uptake values," *J. Nucl. Med.* **46**, 424–428 (2005).
29. R. Boellaard et al., "Effects of noise, image resolution, and ROI definition on the accuracy of standard uptake values: a simulation study," *J. Nucl. Med.* **45**, 1519–1527 (2004).
30. S. Beaulieu et al., "SUV varies with time after injection in 18F-FDG PET of breast cancer: characterization and method to adjust for time differences," *J. Nucl. Med.* **44**, 1044–1050 (2003).
31. F. Crippa et al., "The influence of blood glucose levels on [¹⁸F] fluorodeoxyglucose (FDG) uptake in cancer: a PET study in liver metastases from colorectal carcinomas," *Tumori* **83**, 748–752 (1997).
32. L. A. Pierce et al., "A digital reference object to analyze calculation accuracy of PET standardized uptake value," *Radiology* **277**, 538–545 (2015).
33. R. K. Doot et al., "Biases in multicenter longitudinal PET standardized uptake value measurements," *Transl. Oncol.* **7**, 48–54 (2014).
34. J. S. Scheuermann et al., "Qualification of PET scanners for use in multicenter cancer clinical trials: the American College of Radiology Imaging Network experience," *J. Nucl. Med.* **50**, 1187–1193 (2009).
35. J. J. Sunderland and P. E. Christian, "Quantitative PET/CT scanner performance characterization based upon the society of nuclear medicine and molecular imaging clinical trials network," *J. Nucl. Med.* **56**, 145–152 (2015).
36. M. M. Graham et al., "Summary of the UPICT Protocol for 18F-FDG PET/CT imaging in oncology clinical trials," *J. Nucl. Med.* **56**, 955–961 (2015).
37. R. Boellaard et al., "FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0," *Eur. J. Nucl. Med. Mol. Imaging* **42**, 328–354 (2015).
38. E. Prieto et al., "Evaluation of spatial resolution of a PET scanner through the simulation and experimental measurement of the recovery coefficient," *Comput. Biol. Med.* **40**, 75–80 (2010).
39. M. A. Lodge, A. Rahmim, and R. L. Wahl, "A practical, automated quality assurance method for measuring spatial resolution in PET," *J. Nucl. Med.* **50**, 1307–1314 (2009).
40. D. Byrd et al., "Simultaneous estimation of bias and resolution in PET images with a single long-lived phantom," *J. Nucl. Med.* **55**, 2152–2152 (2014).
41. R. L. Wahl et al., "From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors," *J. Nucl. Med.* **50**, 122S–150S (2009).
42. FDG-PET/CT Biomarker Committee, *FDG-PET/CT as an Imaging Biomarker Measuring Response to Cancer Therapy Profile*, Version 1.13, RSNA.ORG/QIBA (2013).

Darrin Byrd has worked in the Imaging Research Laboratory, the University of Washington since 2011. He received his BA degree in physics from the University of California, Berkeley, and his MS degree in physics from the University of Wisconsin, Milwaukee. He is currently involved in numerous projects related to improving quality control for PET/CT imaging.

Rebecca Christopf received her bachelor's degree in physics and participated in an internship at Stanford Linear Accelerator Center as an undergraduate. Her professional expertise involves management and facilitation of multimodality clinical imaging studies as well as longitudinal observational studies.

John Buatti has been a chair of the Department of Radiation Oncology at the University of Iowa since 2001. In 2005, he opened the Center of Excellence in Image Guided Radiation Therapy, an innovative facility dedicated to quantitative imaging in radiation oncology that includes a 3T MR scanner and 40 slice respiratory-gated PET/CT scanner for simulators to better locate and treat tumors. He is a leading authority on the treatment of cranial malignancies using stereotactic radiosurgery.

Eduardo Moros is a chief of Medical Physics at the Moffitt Cancer Center. He is an active member of the AAPM and ASTRO, an associate editor of the *Journal of Applied Clinical Medical Physics* and the *Journal of Radiation Oncology*. He has published more than 165 peer-reviewed papers and has been a principal investigator/coinvestigator on numerous NIH and industry research grants. His current interest is radiomics applications in radiation oncology.

Sadek Nehmeh is an associate professor of physics in radiology and a chief of PET and nuclear physics at Cornell Medical College, New York, New York, USA. His research interest focuses on the use of PET imaging as a means to monitor response and predict outcome. His work involves imaging of tumor metabolism, hypoxia, and proliferation, as well as prostate 68Ga-PSMA and cardiac 13NH3 imaging.

Adam Opanowski is a senior manager at the Clinical Imaging for Progenics Pharmaceuticals, Inc., located in One World Trade Center, New York, USA. In his current role, he is assisting to develop innovative drugs to improve the lives of prostate cancer patients with drugs in preclinical through late stage development. He has almost 15 years of experience in nuclear medicine and PET imaging. Over the past several years, he has worked with leading imaging researchers on the development of protocol specifications for advanced imaging methods, e.g., kinetic modeling of dynamic PET at the American College of Radiology Imaging Network Center for Research and Innovation. He is also a member of the SNMMI CTN Site Education Committee and is currently the president of the Delaware Valley Society of Nuclear Medicine Technologist.

Paul Kinahan received his BSc and MSc degrees in engineering physics from the University of British Columbia and his PhD in bioengineering in 1994 from the University of Pennsylvania. He is a professor and vice chair for research in the Department of Radiology, University of Washington, with joint appointments in physics, radiation oncology, and bioengineering. He is a director of UWMC PET/CT imaging physics and head of the Imaging Research Laboratory.