SUV variability in EARL-accredited conventional and digital PET

Daniëlle Koopman 1,2 *, Pieter L. Jager 1, Cornelis H. Slump 2, Siert Knollema 1 and Jorn A. van Dalen 3

Abstract

Background: A high SUV-reproducibility is crucial when different PET scanners are in use. We evaluated the SUV variability in whole-body FDG-PET scans of patients with suspected or proven cancer using an EARL-accredited conventional and digital PET scanner.

In a head-to-head comparison we studied images of 50 patients acquired on a conventional scanner (cPET, Ingenuity TF PET/CT, Philips) and compared them with images acquired on a digital scanner (dPET, Vereos PET/CT, Philips). The PET scanning order was randomised and EARL-compatible reconstructions were applied. We measured SUV mean, SUV peak, SUV max and lesion diameter in up to 5 FDG-positive lesions per patient. The relative difference ΔSUV between cPET and dPET was calculated for each SUV-parameter. Furthermore, we calculated repeatability coefficients, reflecting the 95% confidence interval of ΔSUV.

Results: We included 128 lesions with an average size of 19 ± 14 mm. Average ΔSUVs were 6-8% with dPET values being higher for all three SUV-parameters (p < 0.001). ΔSUV max was significantly higher than ΔSUV mean (8% vs. 6%, p = 0.002) and than ΔSUV peak (8% vs. 7%, p = 0.03). Repeatability coefficients across individual lesions were 27% (ΔSUV mean and ΔSUV peak) and 33% (ΔSUV max) (p < 0.001).

Conclusions: With EARL-accredited conventional and digital PET, we found a limited SUV variability with average differences up to 8%. Furthermore, only a limited number of lesions showed a SUV difference of more than 30%. These findings indicate that EARL standardisation works.

Trial registration: This prospective study was registered on the 31th of October 2017 at ClinicalTrials.gov. URL: https://clinicaltrials.gov/ct2/show/NCT03457506?id=03457506&rank=1.

Keywords: FDG-PET, EARL-accreditation, Conventional PET, Digital PET, Cancer

Background

Positron emission tomography/computed tomography (PET/CT) using fluor-18 fluorodeoxyglucose (FDG) is widely used for tumour imaging in patients with cancer. There are ongoing efforts towards standardisation of FDG-PET imaging to allow a quantitative comparison between patients, scanners and medical centres. In 2009 and 2015 the European Association of Nuclear Medicine (EANM) published procedure guidelines on FDG-PET/CT tumour imaging [1, 2]. Furthermore, the EANM launched the EANM Research Ltd. (EARL) to promote nuclear medicine research, including multi-centre trials.

In 2010, EARL started an accreditation program for FDG-PET/CT tumour imaging. This includes EARL-accreditation requirements based on activity concentration recovery coefficients (CRCs) as measured in PET images of a NEMA NU2-2001 image quality phantom. A recent evaluation among the first 200 accredited systems from 150 sites worldwide showed that setting up a harmonising accreditation program is feasible and achievable, and that the FDG-PET/CT program has reduced the variability in semi-quantitative PET performance [3].

Recently, time-of-flight (TOF) PET systems with silicon photomultipliers (SiPM) with digital readout were introduced in clinical practice [4–6]. Although these systems potentially improve image quality compared with PET systems using conventional photomultiplier technology, they can also fulfil EARL accreditation specifications for tumour
imaging with FDG-PET/CT when appropriate reconstruction settings are used [6, 7]. Hence, independent of detector technology, PET systems should provide comparable semi-quantitative results once they fulfill EARL specifications. To our knowledge, this has not yet been explored in clinical practice in a substantial group of patients. Therefore, our aim was to investigate the variability in standardised uptake values (SUVs) on whole-body FDG-PET scans from patients with cancer, using both a conventional and digital EARL-accredited PET scanner.

Materials and methods

Inclusion
We performed a prospective single-centre side-by-side comparison study in 50 patients with suspected or proven cancer who were referred for whole-body FDG-PET/CT. Written informed consent was obtained from all participants included in this study. The Medical Ethical Committee of our institution (METC Isala, Zwolle, Netherlands) approved the study protocol (NL52329.075.15).

PET/CT acquisition
 Patients fasted for at least 6 h prior to the PET scan. Blood glucose levels were measured before intravenous injection of FDG, to ensure a value below 10 mmol/L. Patients were administered a FDG-activity based on \( A = 6.2 \ \text{w}^2/t \), where \( A \) is the FDG-activity administered in Megabecquerel (MBq), \( w \) is the patient’s body weight in kilogram (kg) and \( t \) is the acquisition time per bed position in seconds (s) [8].

For each patient whole-body PET scans from head to groin were acquired in supine position using a state-of-the-art TOF PET/CT scanner with conventional photomultiplier technology (cPET, Ingenuity TF, Philips Healthcare) and a TOF PET/CT scanner with digital SIPMs and digital readout (dPET, Vereos, Philips Healthcare). Both systems were EARL-accredited. For both PET scanners the error in cross-calibration with the associated dose calibrator was less than 5%. The PET scanning order was randomised per patient. We included 25 patients who were first scanned on dPET and afterwards on cPET (dPET-first group), and we included 25 patients who were first scanned on cPET and afterwards on dPET (dPET-second group). Per patient and per scan we collected \( \Delta t \) which was defined as the time between FDG-administration and the start of the PET scan.

PET acquisition times of the first scan were 72 s and 144 s per bed position for patients with body weight ≤ 80 kg and > 80 kg, respectively. For the second scan the scan time per bed position was equal to the scan time of the first scan plus a compensation for the radioactive decay of fluor-18. The resulting average scan time of the second PET scan was 85 s (range 72–91 s) for patients ≤ 80 kg and 180 s (range 147–205 s) for patients > 80 kg.

Prior to each PET scan a CT scan was acquired for attenuation correction. The CT scan parameters were 120 kV, 64 mAs (range 39–136 mAs), 64 × 0.625 mm slice collimation, a pitch of 0.83 and a rotation time of 0.5 s.

PET/CT reconstruction
For both systems we used EARL-compatible reconstructions. For cPET an ordered subset expectation maximisation (OSEM) TOF PET reconstruction was applied with \( 4 \times 4 \times 4 \) mm\(^3\) voxels and a relaxation parameter of 1.0, without point spread function (PSF) modelling, as previously described [9]. For dPET we performed an OSEM TOF PET reconstruction with \( 4 \times 4 \times 4 \) mm\(^3\) voxels and a 3-mm Gaussian post-smoothing filter, without PSF modelling, as previously described [7]. For both cPET and dPET attenuation correction was applied using iteratively reconstructed CT data with iDose level 4 and a slice thickness of 3 mm.

Semi-quantitative evaluation
Semi-quantitative analyses were performed using the quAntitative onCology moleCUlar Analysis suiTE (ACCURATE) tool [10]. For each patient we included a maximum of 5 FDG-positive lesions, to prevent a possible bias from patients with many lesions. In case a patient had more than 5 eligible lesions, we selected the 5 lesions with the shortest diameter on the CT scan and which were measurable on both PET scans using the ACCURATE tool. We chose this selection approach because smaller lesions can be more sensitive to reconstruction differences.

For each lesion we measured the mean, peak and maximum standardised uptake value (SUV\(_{\text{mean}}\), SUV\(_{\text{peak}}\) and SUV\(_{\text{max}}\)) on cPET and dPET images. SUV\(_{\text{mean}}\) was based on the 3D isocontour derived at 50% of the maximum pixel value. SUV\(_{\text{peak}}\) was defined as the average SUV of a spherical 1 cm\(^3\) volume-of-interest in the tumour-region with the highest uptake [11]. Furthermore, we measured the short-axis diameter on the axial slice of the CT scan.

Following the paper by Lodge [12] we calculated the relative difference \( \Delta \text{SUV} \) per lesion between cPET and dPET for SUV\(_{\text{mean}}\), SUV\(_{\text{peak}}\) and SUV\(_{\text{max}}\) using Eq. 1.

\[
\Delta \text{SUV} = \frac{\text{SUV}_{\text{dPET}} - \text{SUV}_{\text{cPET}}}{(\text{SUV}_{\text{dPET}} + \text{SUV}_{\text{cPET}}) \times 0.5}
\]

In addition, we derived the standard deviation (SD) of \( \Delta \text{SUV} \) and we calculated the repeatability coefficient (RC) using Eq. 2.

\[
\text{RC} = 1.96 \times SD(\Delta \text{SUV})
\]
cut-off value is considered by PERCIST to indicate a switch from “stable” disease to either “progression” or “response” [13].

Statistical analysis
The statistical analysis was performed using SPSS Version 24. Quantitative results were presented as mean ± SD. Data distribution normality was evaluated using the Shapiro-Wilk test. For data that were not normally distributed the median was included as well. We performed an independent-sample t test to compare patient and scan characteristics (age, body weight, administered FDG-activity and ΔT) between patients in both scanning groups. Furthermore, we performed the Mann-Whitney U-test to compare lesion diameters between lesions in both scanning groups. Differences in average SUVmean, SUVpeak and SUVmax between cPET and dPET were evaluated with the Wilcoxon signed-rank test. To test whether average ΔSUV differences between the two PET systems were similar for the three SUV-parameters, we pairwise compared ΔSUVmean, ΔSUVpeak and ΔSUVmax using a paired 2-sample t test. Furthermore, we performed the Pitman-Morgan test (using R studio, package PairedData) to pairwise compare the RCs of the three SUV-parameters. Moreover, we performed a linear regression analysis (Pearson’s correlation coefficient and F-test) to determine correlations between ΔSUV and the time between FDG-administration and the start of the dPET scan (ΔTdPET), and between ΔSUV and lesion diameter. A p value less than 0.05 was considered to indicate statistical significance.

Results
Patient characteristics
We included 50 patients (27 males, 23 females) with suspected or proven lung cancer (n = 35), breast cancer (n = 8), lymphoma (n = 3), oesophageal cancer (n = 3) or gastric cancer (n = 1). Patient and scan characteristics per scanning group are presented in Table 1. The characteristics of both groups were comparable (p ≥ 0.16). In total we evaluated 128 FDG-positive lesions, among which 66 lesions were part of the dPET-first group and 62 lesions of the dPET-second group. The average lesion diameter was 19 ± 14 mm (median 15 mm, range 4–90 mm) with comparable sizes across both scanning groups (p = 0.36). The number of included lesions per patient was 1 in 17 patients, 2 in 11 patients, 3 in 7 patients, 4 in 7 patients and 5 in 8 patients.

Semi-quantitative evaluation
SUVmean, SUVpeak and SUVmax over all 128 lesions are shown in Table 2 and Fig. 1 for cPET and dPET separately. Average dPET values were higher than cPET values for all three SUV-parameters (p < 0.001).

Furthermore, relative SUV differences (ΔSUV) between cPET and dPET are shown in Table 2 and Fig. 2. The average variability in SUVmax was significantly higher than in SUVmean (8% vs. 6%, p = 0.002) and in SUVpeak (8% vs. 7%, p = 0.03), while ΔSUVmean and ΔSUVpeak were similar across all lesions (6% vs. 7%, p = 0.08). Furthermore, corresponding RCs were 27% (ΔSUVmean and ΔSUVpeak) and 33% (ΔSUVmax), with the RC of SUVmax being higher than the RCs of SUVmean and SUVpeak (p < 0.001). SUVmean and SUVpeak RCs were similar (p = 0.35).

The number of lesions with an absolute ΔSUV ≥ 30% was 3 (2%) for SUVmean, 4 (3%) for SUVpeak and 15 (12%) for SUVmax. All lesions but one with a ΔSUV variability of ≥ 30% were part of the dPET-second group.

Correlations between ΔSUV and ΔTdPET are presented in Fig. 3 for all three SUV-parameters. It shows that ΔSUVmean, ΔSUVpeak and ΔSUVmax increased at prolonged ΔTdPET (p < 0.001) with correlation coefficients of 0.54, 0.55 and 0.59, respectively. Furthermore, the average ΔSUV of lesions in the dPET-second group was significantly higher as compared with lesions in the dPET-first group, with ΔSUVmean of 16% and – 3%, respectively (p < 0.001), ΔSUVpeak of 16% and – 2%, respectively (p < 0.001), and ΔSUVmax of 21% and – 4%, respectively (p < 0.001). In Fig. 4 we compared ΔSUV for each lesion with its diameter. We found no correlation between these two parameters (R < 0.09, p > 0.33).

Table 1 Patient (n = 50) and scan characteristics

|                          | dPET-first group (n = 25) | dPET-second group (n = 25) | p value |
|--------------------------|--------------------------|----------------------------|---------|
| Age (in years)           | 64 ± 10                  | 67 ± 12                    | 0.23    |
| Body weight (in kg)      | 83 ± 19                  | 76 ± 16                    | 0.16    |
| Glucose level (in mmol/L)| 5.7 ± 0.8                | 6.0 ± 1.0                  | 0.25    |
| Administered FDG-activity (in MBq) | 413 ± 105 | 397 ± 97                   | 0.60    |
| ΔT from FDG administration |                          |                            |         |
| Until first PET scan (in min) | 64 ± 10 | 66 ± 10                    | 0.51    |
| Until second PET scan (in min) | 96 ± 11 | 97 ± 13                    | 0.79    |

*Continuous variables are described as mean ± SD
Clinical example
In Fig. 5, FDG-PET/CT images are shown from a patient with suspected lung cancer in the dPET-second group. Both PET scans showed bilateral adrenal gland metastases with higher SUVs ($\Delta$SUV 7–15%) on the second dPET scan that was acquired 24 min after the cPET scan.

Discussion
We evaluated the SUV variability in whole-body FDG-PET scans from 50 patients with cancer by comparing conventional and digital EARL-accredited PET. The average SUV variability across 128 FDG-positive lesions was limited with $\Delta$SUVs of 6–8%. Furthermore, only a limited number of lesions showed a SUV difference of more than 30%. These findings indicate that EARL standardisation works.

We compared the variability of three SUV-parameters in a pairwise fashion, and as expected we found the variability in SUV$\text{max}$ to be higher than in SUV$\text{mean}$ and SUV$\text{peak}$ ($p \leq 0.03$), although the average differences were relatively small (8% vs. 6–7%). We used automated software to identify the tumour region with the highest uptake within the lesion, and it has been suggested that this method provides a lower variability for SUV$\text{peak}$ as compared with SUV$\text{max}$ [12]. Recently, EARL adopted SUV$\text{peak}$ as an additional metric in the updated EARL accreditation standards [14], as it appeared to be less sensitive to changes in reconstruction parameters and acquisition durations than SUV$\text{mean}$ or SUV$\text{max}$ [15]. However, a drawback of common SUV$\text{peak}$ definitions is that its volume of 1 cm$^3$ is not suitable for (sub)centimeter lesions [15].

We found repeatability coefficients of 27% (SUV$\text{mean}$ and SUV$\text{peak}$) and 33% (SUV$\text{max}$). This variability is likely caused by a combination of three factors: a difference in

### Table 2

|        | cPET$^a$ | dPET$^a$ | $\Delta$SUV (%)$^a$ | RC  | $p$ value |
|--------|----------|----------|---------------------|-----|-----------|
| SUV$\text{mean}$ | 5.3 ± 3.8 (4.1) | 5.6 ± 4.3 (4.6) | 6% ± 14% | 27% | < 0.001   |
| SUV$\text{peak}$  | 6.4 ± 5.2 (4.7) | 6.8 ± 5.9 (5.2) | 7% ± 14% | 27% | < 0.001   |
| SUV$\text{max}$   | 8.4 ± 6.3 (6.6) | 9.1 ± 7.0 (7.3) | 8% ± 17% | 33% | < 0.001   |

$^a$Continuous variables are described as mean ± SD (and median if not normally distributed)

![Fig. 1](image1.png)

**Fig. 1** SUV$\text{mean}$ (a), SUV$\text{peak}$ (b) and SUV$\text{max}$ (c) as measured on cPET and dPET across all lesions ($n=128$). The y-axis is shown on a log scale. Average dPET values were significantly higher than cPET values for all three parameters ($p < 0.001$). This boxplot shows the median, interquartile range and outliers (o): values that are between 1.5 and 3.0 box length from the percentile borders.

![Fig. 2](image2.png)

**Fig. 2** $\Delta$SUV variability for SUV$\text{mean}$, SUV$\text{peak}$ and SUV$\text{max}$ between cPET and dPET across all lesions ($n=128$). The average variability in $\Delta$SUV$\text{max}$ was larger than the variability in $\Delta$SUV$\text{mean}$ ($p = 0.002$) and $\Delta$SUV$\text{peak}$ ($p = 0.03$). Furthermore, $\Delta$SUV$\text{real}$ had a higher variance as compared with $\Delta$SUV$\text{mean}$ and $\Delta$SUV$\text{peak}$ ($p < 0.001$). This boxplot shows the median and the interquartile range.
EARL CRCs between our cPET and dPET system, the impact of prolonged FDG-uptake and the SUV test-retest variability. These 3 factors are discussed in the next 3 paragraphs.

Concerning CRC differences, the EARL protocol for our dPET system was based on relatively high CRCs for sub-15 mm small spheres [7] as compared with the CRCs of our cPET EARL protocol [9], with 10–20% higher CRCs on dPET. This explains why we found average ΔSUVs of 6–8% with dPET SUVs being higher than cPET values (p < 0.001) in most cases. Larger variations can be expected at other PET sites or in clinical trials that use multiple EARL-accredited PET systems with divergent CRCs. This is possible because current EARL accreditation specifications [16] accept relatively large differences in CRCs, especially for small spheres (Table 3). To further harmonise the semiquantitative results of EARL-accredited PET scanners, PET reconstruction settings could be further aligned to provide more similar CRCs. Naturally, SUV variability could also be reduced by using the same system and therefore this should be applied in longitudinal PET comparisons of the same patient [17].

Concerning the time-interval between the first and the second scan, it is known that SUVs generally increase with prolonged FDG-uptake [18, 19]. We corrected for this effect by randomising the PET scanning order. Consequently, the average ΔSUV across all lesions is likely not influenced by this effect. However, ΔSUVs of individual lesions were higher after the longer interval as

![Fig. 3 Scatterplot comparing the relative change in SUV<sub>mean</sub> (a), SUV<sub>peak</sub> (b) and SUV<sub>max</sub> (c) with ΔTdPET, defined as the time between FDG-administration and start of the dPET scan. ΔSUV<sub>mean</sub>, ΔSUV<sub>peak</sub> and ΔSUV<sub>max</sub> increased with prolonged ΔTdPET (p < 0.001).](image)

![Fig. 4 Scatterplot comparing the relative change in SUV<sub>mean</sub> (a), SUV<sub>peak</sub> (b) and SUV<sub>max</sub> (c) with lesion diameter. The x-axis is shown on a log scale. There were no significant correlations between ΔSUV and lesion diameter with R = 0.09 for ΔSUV<sub>mean</sub> and ΔSUV<sub>peak</sub> (p = 0.32), and R = 0.01 for ΔSUV<sub>max</sub> (p = 0.96).](image)
shown in Fig. 3. It is likely that the higher average ΔSUV for lesions in the dPET-second group is both caused by an increase in SUV due to prolonged FDG-uptake as well as the effect of the digital scanner with its higher CRCs. Conversely, in the dPET-first group an increase in SUV on the second scan caused by prolonged FDG-uptake is partly being compensated as cPET images were based on a reconstruction with lower CRCs as compared with dPET. For example, the average ΔSUVmax in the dPET-first group was −4% whereas the average ΔSUV-max in dPET-second group was 21%. Based on these averages, we expect that about (21% + 4%)/2 = 13% of the higher SUV max on the second scan can be attributed to the prolonged FDG-uptake. If this theoretical correction of 13% is applied to all individual ΔSUVs, only 1 lesion (1%) remains with a ΔSUVmax ≥30%.

Concerning the SUV test-retest variability, it is known that biological, technical and methodological factors [12, 19] play a role. Several studies have evaluated this in different types of cancer [12, 20–22]. In a recent review, Lodge [12] stated that with a strict protocol, lesion-SUV has a within-subject coefficient of variation (wCV) of 10% (SUV_mean and SUV_peak) and 11% (SUV_max). In our study, we found RCs of 27–33%, representing wCVs of 10% (SUV_mean and SUV_peak) and 12% (SUV_max) when using wCV = RC/(√2 × 1.96). This indicates that the average ΔSUV in our study is comparable with values reported by Lodge [12]. However, our study includes two aspects that make it difficult to compare our wCVs with the data reported by Lodge. First, we performed both PET scans on the same day after a single FDG-administration while Lodge [12] only included results based on two separate FDG-administrations. Second, the lesions that we included were relatively small (median size 15 mm) while Lodge [12] described that most repeatability studies included lesions with a minimum diameter of 20 to 30 mm.

Our conclusion that EARL standardisation works is in agreement with findings from a recently published paper by van Sluis et al. [23]. They performed a cPET versus dPET comparison study, using scanners from another vendor, in a small group of patients with cancer (n = 20). Although they did not calculate relative differences or repeatability coefficients, they observed a good agreement in SUV measurements between both PET/CT systems, in particular when using EARL-compliant reconstructions on both systems [23].

The present study has some limitations. We included 128 lesions across 50 patients where the included number of lesions varied between 1 and 5 lesions per patient, but we did not take a possible intra-patient correlation between lesions into account in the statistical analysis. Yet, the number of lesions in both scanning groups was almost similar (66 vs. 62 lesions). Furthermore, our

![Fig. 5 Axial FDG-PET/CT images (a, b and c) from a lung cancer patient with bilateral adrenal gland metastases showing higher SUVs on the dPET scan (b) that was acquired 24 min after the cPET scan (a). The left-gland metastasis (diameter 11 mm) showed ΔSUVs of 7% (SUV_mean), 10% (SUV_max) and 15% (SUV_peak). ΔSUVs of the right-gland metastasis (diameter 14 mm) were 13% (SUV_mean) and 11% (SUV_peak and SUV_max). In this case the impact of the digital scanner cannot be separated from the SUV rise caused by the prolonged FDG-uptake. Meanwhile, the visual image quality of both PET scans appeared comparable in terms of image noise, texture and FDG-uptake as intended with an EARL-compatible protocol.]

| Sphere diameter | Limits | CRC_min–max limits | CRC_min–max | CRC_mean (%) | CRC_mean (%) | CRC_min–max limits | CRC_min–max | CRC_max (%) | CRC_max (%) |
|-----------------|--------|-------------------|-------------|-------------|-------------|-------------------|-------------|-------------|-------------|
| 10 mm           | 0.27–0.43 | 46%               | 0.34–0.57   | 51%         |             |                   |             |             |             |
| 13 mm           | 0.44–0.60 | 31%               | 0.59–0.85   | 36%         |             |                   |             |             |             |
| 17 mm           | 0.57–0.73 | 25%               | 0.73–1.01   | 32%         |             |                   |             |             |             |
| 22 mm           | 0.63–0.78 | 21%               | 0.83–1.09   | 27%         |             |                   |             |             |             |
| 28 mm           | 0.72–0.85 | 17%               | 0.91–1.13   | 22%         |             |                   |             |             |             |
| 37 mm           | 0.76–0.89 | 16%               | 0.95–1.16   | 20%         |             |                   |             |             |             |
study was not a full test-retest study since for each patient both PET scans were acquired on the same day and with just a single FDG-injection. Therefore, variability associated with patient preparation, biological factors and FDG-administration was not fully taken into account in our study. However, other factors such as patient motion, breathing and potential CT-PET mismatches could still have influenced the ΔSUV in this intra-individual comparison of EARL-accredited PET and dPET scans. Still, given that the impacts of the PET systems, biological effect and test-retest are intricate and that biological effects are not negligible, it would be useful to repeat this semi-quantitative comparison of EARL-accredited PET scans in a full test-retest setting to confirm our results. Another limitation is the wide range in ΔT for the second scan as shown in Fig. 3, which influences individual ΔSUVs. Fortunately, the average FDG-uptake time per scan between both scanning groups was similar.

While the present study is based on current EARL accreditation specifications [16], an update of those specifications has been proposed because in recent years different vendors launched new PET/CT systems equipped with novel techniques such as TOF, resolution modelling/PSF technologies and digital detectors. These modern systems can deliver PET images with higher CRCs, especially for small spheres, and therefore, an update of the EARL accreditation specifications is desirable. Kaalep et al. [15] evaluated the feasibility of harmonising performance for novel PET/CT systems, and they also proposed new EARL criteria. In these newly proposed CRCs the relative difference (%) between upper and lower limits is similar to current EARL specifications [16]. Therefore, it is expected that the potential variability in semi-quantitative FDG-PET with such updated EARL-compatible protocols will remain similar.

**Conclusion**

With EARL-accredited conventional and digital PET, we found a limited SUV variability with average differences up to 8%. Furthermore, only a limited number of lesions showed a SUV difference of more than 30%. These findings indicate that EARL standardisation works. When EARL-accredited systems with divergent CRCs are used, larger SUV differences can be expected.

**Abbreviations**

ΔT: Time between FDG-administration and start of the PET scan; CRC: Contrast recovery coefficient; CT: Computed tomography; EANM: European Association of Nuclear Medicine; EARL: EANM Research Ltd; FDG: Fluor-18 fluorodeoxyglucose; MIP: Maximum intensity projection; OSEM: Ordered subset expectation maximisation; PET: Positron emission tomography; PSF: Point spread function; RC: Repeatability coefficient; SD: Standard deviation; SUV: Standardised uptake value; SUV max: Maximum SUV; SUV mean: Average SUV; SUV peak: Peak SUV; TOF: Time-of-flight; \( \nu_{CV} \): Within-subject coefficient of variation

**Acknowledgements**

We gratefully acknowledge the staff from Isala, Zwolle and Technical Medicine Center, Enschede, for their overall support and kind collaboration.

**Authors’ information**

Not applicable.

**Authors’ contributions**

DK and JvD designed the study. DK performed the quantitative PET evaluation and analysed the data. DK wrote the concept manuscript that was revised by JvD, CS, SK and PJ. All authors approved the final manuscript.

**Funding**

This work was supported by a research agreement between the Department of Nuclear Medicine, Isala, and Philips Healthcare regarding new PET technologies.

**Availability of data and materials**

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

**Ethics approval and consent to participate**

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This research involved human participants. We received approval from the Medical Ethical Committee of our institution (Isala, Zwolle) to perform this study (NL52329.075.15). Informed consent was obtained from all individual participants included in the study.

**Consent for publication**

Not applicable.

**Competing interests**

This work was supported by a research agreement between the Department of Nuclear Medicine, Isala, and Philips Healthcare regarding new PET technologies. The content of the article was solely the responsibility of the authors. The authors declare that they have no competing interests.

**Author details**

1Department of Nuclear Medicine, Isala, Dokter van Heesweg 2, 8025, AB, Zwolle, the Netherlands. 2Technical Medicine Center, University of Twente, Enschede, the Netherlands. 3Department of Medical Physics, Isala, Zwolle, the Netherlands.

Received: 11 September 2019 Accepted: 16 October 2019
Published online: 10 December 2019

**References**

1. Boellaard R, O’Doherty MJ, Weber WA, Mottaghy FM, Lonsdale MN, Snoebants SG, WKG O, Kotzerke J, Hoekstra OS, Prum J. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. Eur J Nucl Med Mol Imaging. 2010;37:181–200.
2. Boellaard R, Delgado-Bolton R, Oyen WJG, Giamparle F, Tatsch K, Eschner W, Verzijlbergen FJ, Barrington SF, Pike LC, Weber WA. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. Eur J Nucl Med Mol Imaging. 2015;42:328–54.
3. Kaalep A, Sera T, Oyen WJG, Krause BJ, Chiti A, Liu Y, Boellaard R, EANM/EARL FDG-PET/CT accreditation – summary results from the first 200 accredited imaging systems. Eur J Nucl Med Mol Imaging. 2018;45:411–22.
4. Miller M, Zhang J, Binzel K, Griesmer J, Laurence T, Narayanan M, Natarajamani D, Wang S, Knopp M. Characterization of the varoius digital photon counting PET system. J Nucl Med. 2015;56:434.
5. Hsu DF, Illan E, Peterson WT, Uribe J, Lubberink M, Levin CS. Studies of a next-generation silicon-photomultiplier-based time-of-flight PET/CT system. J Nucl Med. 2017;58:1511–8.
6. van Sluis JJ, de Jong J, Schar J, Noordzij W, van Snick P, Dieckx R, Borra R, Willemsen A, Boellaard R. Performance characteristics of the digital Biograph Vision PET/CT system. J Nucl Med. 2019;60:1031–6.
7. Koopman D, Koerkamp MG, Jager PL, Arkies H, Knollema S, Slump CH, Sanches PG, van Dalen JA. Digital PET compliance to EARL accreditation specifications. EJNMMI Phys. 2017;4:9–14.
8. de Groot EH, Post N, Boellaard R, NRL W, ATM W, van Dalen JA. Optimized dose regimen for whole-body FDG-PET imaging. EJNMMI Res. 2013;3:63–72.
9. Koopman D, van Dalen IA, MCM L, Arkies H, de Boer J, AHJ O, Slump CH, Jager PL. Improving the detection of small lesions using a state-of-the-art time-of-flight PET/CT system and small-voxel reconstructions. J Nucl Med Technol. 2015;43:21–7.
10. Boellaard R. Quantitative oncology molecular analysis suite: ACCURATE. J Nucl Med. 2018;59:1753.
11. Sher A, Lacoeuille F, Fosse P, Venueren L, Cahuet-Vannier A, Dabli D, Bouchet F, Couturier O. For avid glucose tumors, the SUV peak is the most reliable parameter for [18 F] FDG-PET/CT quantification, regardless of acquisition time. EJNMMI Res. 2016;6:21–6.
12. Lodge MA. Repeatability of SUV in Oncologic 18F-FDG PET. J Nucl Med. 2017;58:523–32.
13. Wahl RL, Jacene H, Kasamon Y, Lodge MA. From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors. J Nucl Med. 2009;50:1225.
14. EARL. Updated EANM FDG PET/CT accreditation specifications for SUV recovery coefficients [2019]. 2019. http://earl.eanm.org/cms/website.php. Accessed 01-08- 2019.
15. Kaalep A, Sera T, Rijnsdorp S, Yaqub M, Talsma A, Lodge MA, Boellaard R. Feasibility of state of the art PET/CT systems performance harmonisation. Eur J Nucl Med Mol Imaging. 2018;45:1344–61.
16. EARL. EANM FDG PET/CT accreditation specifications for SUV recovery coefficients. 2017. http://earl.eanm.org/cms/website.php . Accessed 01-10- 2018.
17. Boellaard R, Delgado-Bolton R, Oyen WJG, Giammarile F, Tatsch K, Eschner W, Verrijldben FJ, Barrington SF, Pike LC, Weber WA. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. Eur J Nucl Med Mol Imaging. 2015;42:328–54.
18. Shankar LK, Hoffman JM, Bacharach S, Graham MM, Karp J, Lammertsma AA, Larson S, Mankoff DA, Siegel BA, Van den Abbeelee A. Consensus recommendations for the use of 18F-FDG PET as an indicator of therapeutic response in patients in National Cancer Institute Trials. J Nucl Med. 2006;47:1059–66.
19. Vriens D, Visser EP, de Geus-Oei L, Oyen WJG. Methodological considerations in quantification of oncological FDG PET studies. Eur J Nucl Med Mol Imaging. 2010;37:1408–25.
20. de Langen AJ, Vincent A, Velasquez LM, van Tinteren H, Boellaard R, Shankar LK, Boer M, Smit EF, Stroobants S, Weber WA. Repeatability of 18F-FDG uptake measurements in tumors: a metaanalysis. J Nucl Med. 2012;53:701.
21. Rockall AG, Avril N, Lam R, Iannone R, Mozley PD, Parkinson C, Bergstrom DA, Sala E, Sarker S, McNeish IA. Repeatability of quantitative FDG-PET/CT and contrast enhanced CT in recurrent ovarian carcinoma: test retest measurements for tumor FDG uptake, diameter and volume. Clin Cancer Res. 2014;20:2751–60.
22. Weber WA, Gatsonis CA, Mozley PD, Hanna LG, Shields AF, Abere DR, Govindan R, Torligni DA, Karp JS, Yu JQM. Repeatability of 18F-FDG PET/CT in advanced non–small cell lung cancer: prospective assessment in 2 multicenter trials. J Nucl Med. 2015;56:1137–43.
23. van Sluis J, Boellaard R, Somasundaram A, van Snick P, Borra R, Dierckx R, Stormezand G, Glaudemans A, Noordzij W. Image quality and semi-quantitative measurements of the Siemens Biograph Vision PET/CT: initial experiences and comparison with Siemens Biograph mCT PET/CT. J Nucl Med. 2019.

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