Received:
3 October 2014

Revised:
21 January 2015

Accepted:
29 January 2015

doi: 10.1259/bjr.20140655

Cite this article as:

Rasmussen JH, Fischer BM, Aznar MC, Hansen AE, Vogelius IR, Löfgren J, et al. Reproducibility of ^{18}F -FDG PET uptake measurements in head and neck squamous cell carcinoma on both PET/CT and PET/MR. Br J Radiol 2015;88:20140655.

FULL PAPER

Reproducibility of ^{18}F -FDG PET uptake measurements in head and neck squamous cell carcinoma on both PET/CT and PET/MR

¹J H RASMUSSEN, MD, ²B M FISCHER, MD, DMSc, ¹M C AZNAR, PhD, ²A E HANSEN, PhD, ¹I R VOGELIUS, PhD, ²J LÖFGREN, MD, ²F L ANDERSEN, PhD, ²A LOFT, MD, PhD, ²A KJAER, MD, DMSc, ²L HØJGAARD, MD, DMSc and ¹L SPECHT, MD, DMSc

¹Department of Oncology, Section of Radiotherapy, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark

²Department of Clinical Physiology, Nuclear Medicine & PET, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark

Address correspondence to: Dr Jacob Høygaard Rasmussen
E-mail: jacobrasmu@gmail.com

Objective: To investigate reproducibility of fluorine-18 fludeoxyglucose (^{18}F -FDG) uptake on ^{18}F -FDG positron emission tomography (PET)/CT and ^{18}F -FDG PET/MR scans in patients with head and neck squamous cell carcinoma (HNSCC).

Methods: 30 patients with HNSCC were included in this prospective study. The patients were scanned twice before radiotherapy treatment with both PET/CT and PET/MR. Patients were scanned on the same scanners, 3 days apart and according to the same protocol. Metabolic tumour activity was measured by the maximum and peak standardized uptake value (SUV_{max} and SUV_{peak} , respectively), and total lesion glycolysis from the metabolic tumour volume defined from $\geq 50\%$ SUV_{max} . Bland-Altman analysis with limits of agreement, coefficient of variation (CV) from the two modalities were performed in order to test the reproducibility. Furthermore, CVs from SUV_{max} and SUV_{peak}

were compared. The area under the curve from cumulative SUV-volume histograms were measured and tested for reproducibility of the distribution of ^{18}F -FDG uptake.

Results: 24 patients had two pre-treatment PET/CT scans and 21 patients had two pre-treatment PET/MR scans available for further analyses. Mean difference for SUV_{max} , peak and mean was approximately 4% for PET/CT and 3% for PET/MR, with 95% limits of agreement less than $\pm 20\%$. CV was small (5–7%) for both modalities. There was no significant difference in CVs between PET/CT and PET/MR ($p = 0.31$). SUV_{max} was not more reproducible than SUV_{peak} ($p = 0.09$).

Conclusion: ^{18}F -FDG uptake in PET/CT and PET/MR is highly reproducible and we found no difference in reproducibility between PET/CT and PET/MR.

Advances in knowledge: This is the first report to test reproducibility of PET/CT and PET/MR.

Functional imaging with fluorine-18 fludeoxyglucose positron emission tomography combined with CT (^{18}F -FDG PET/CT) has been shown to be useful for prognostication of head and neck squamous cell carcinoma (HNSCC),^{1–3} and the use of ^{18}F -FDG PET/CT has also been shown to reduce interobserver variability in target delineation for radiotherapy.^{4,5} Furthermore, ^{18}F -FDG PET/CT can identify regions of the tumour with a high risk of relapse, leading to the idea that ^{18}F -FDG uptake might be a target for dose painting.^{6,7} Finally, ^{18}F -FDG PET/CT may be used in response evaluation.^{8,9} Maximum standardized uptake value (SUV_{max}) has for many years been the main uptake measurement in prognostic studies for various malignancies. More recent studies have focused on demonstrating prognostic value of PET/CT-based volumetric parameters such as metabolic tumour volume (MTV) and total lesion glycolysis (TLG). MTV is the sum of the volume of voxels with standard uptake value (SUV) exceeding a certain

threshold value in a tumour, and TLG is calculated by multiplying MTV and the mean standardized uptake value (SUV_{mean}) of the MTV. These volume-based PET parameters have increasingly gained interest and have been reported to be significant prognostic factors for various malignancies including HNSCC.^{10–13} ^{18}F -FDG PET/CT is currently not routinely recommended as a diagnostic tool in HNSCC except in very specific situations,¹⁴ but reproducibility of the ^{18}F -FDG signal is a prerequisite for a more widespread use of ^{18}F -FDG PET for the above-mentioned indications. Yet, only a few studies of the reproducibility of ^{18}F -FDG PET/CT exist^{8,15–21} and none of these studies includes patients with HNSCC.

MRI is gaining acceptance as an imaging modality for oncology as it offers superior soft-tissue contrast compared with CT alone, and it has been suggested that information from PET/CT and MR is complementary in head and neck

cancer.²² The introduction of the integrated PET/MR scanner offers a unique opportunity to combine the high soft-tissue contrast of MR with the functional imaging from PET within a single imaging session. PET/MR is still in its infancy, but the combined modality imaging is potentially useful in the management of patients with HNSCC.^{22–28} However, the same criteria of reproducibility as with PET/CT should be upheld by this new modality. The purpose of this prospective test–retest study is to assess the reproducibility of both ¹⁸F-FDG PET/CT and ¹⁸F-FDG PET/MR in a homogenous cohort of patients with HNSCC.

METHODS AND MATERIALS

Patients

30 patients with HNSCC in the pharynx referred to radiotherapy were included in this prospective study. Patients with diabetes were excluded. All patients gave informed consent, and the study was approved by the local ethics committee, approval number H-3-2012-072. The patients were instructed to fast for at least 6 h before the examination. All patients were scanned prior to treatment with repeated scans on both modalities, exactly 3 days apart.

Fluorine-18 fludeoxyglucose positron emission tomography/CT

All PET/CT scans were performed on the same biograph mCT 64 (Siemens Medical Solutions, Malvern, PA), and patients were immobilized with an individually moulded thermoplastic mask in the supine position on a flat scanner couch. On Day 1, the scan was a whole-body PET/CT (apex of the skull to mid-thigh), and on Day 2, the scan covered the apex of the skull to apex of the lung. All CT images were acquired as diagnostic quality spiral CTs using 120 kVp and intravenous contrast agent, except for one patient who had renal insufficiency. A multibed PET scan with 2 min per bed position was performed after the CT, and the patients were scanned 60 min after ¹⁸F-FDG injection (4 MBq kg^{−1}). In case of a delay on Day 1, a match of the actual time from ¹⁸F-FDG injection to scan was attempted on the second scanning day.

PET/CT images were reconstructed using an iterative three-dimensional ordinary poisson-ordered subset expectation maximization (OP-OSEM) algorithm using resolution modelling of the point spread function (PSF) and time of flight (TOF) with 2-mm gaussian post filter, which theoretically provides superior image reconstruction^{29–31} and is used in our clinical routine. However, on PET/MR, TOF is not available and PSF not comparable. To improve cross-modality comparison, an additional reconstruction of PET/CT data was carried out without PSF and TOF and with 4-mm gaussian post-filter. The two sets of PET data were denoted as “PET/CT (PSF/TOF)” and “PET/CT (OSEM)”, respectively. For both PET reconstructions, 3 iterations, 21 subsets and 400 × 400 matrices (2.0 × 2.0-mm voxels) were used.

Fluorine-18 fludeoxyglucose positron emission tomography/MR

The patients were scanned with an integrated PET/MR system (Siemens Biograph mMR) with a 3-T magnet using a head and neck coil. The MR sequences acquired were *T*₂ weighted as well

as a Dixon VIBE sequence for attenuation correction purposes. PET was performed as a single-bed, 20-min acquisition. The patients were planned to be scanned 100–120 min after ¹⁸F-FDG injection (after the PET/CT scan). A match of the ¹⁸F-FDG injection to PET/MR time interval on the second scanning day to the actual time interval on the first scanning day was attempted. PET data were reconstructed using OP-OSEM with 3 iterations, 24 subsets and 4-mm gaussian post-filter into 344 × 344 matrices to match the PET/CT reconstructions as close as possible. Resolution modelling (PSF) was not applied.

Imaging analysis and measurements of standardized uptake value

Lesions were counted from PET/CT (PSF/TOF) and PET/MR to make a simple test–retest measurement of the two modalities. The region of interest was defined as oropharynx, and in one case hypopharynx, including neck Levels I–V in case of node metastasis. An isocontour corresponding to 50% of SUV_{max} was generated and subsequently adapted by excluding regions with physiological ¹⁸F-FDG uptake by an experienced nuclear

Table 1. Number of scans and patient characteristics of the 24 included patients

PET/CT	PET/MR	Number of patients (%)
Number of scans		
2	2	24
2	1	3
1	2	0
1	1	3
Patient characteristics of the 24 included patients		
Age (years)		
Median		64
Range		44–74
Sex		
Male		20 (83.3)
Female		4 (16.7)
T-stage		
1		0 (0)
2		17 (70.8)
3		7 (29.2)
4		0 (0)
N-stage		
0		9 (37.5)
1		2 (8.3)
2a		1 (4.2)
2b		5 (20.8)
2c		6 (25.0)
3		1 (4.2)

PET, positron emission tomography.

Figure 1. Flowchart illustrating patient inclusion and exclusion. ^{18}F -FDG, fluorine-18 fludeoxyglucose; PET, positron emission tomography.

medicine physician (BMF). Only one MTV per patient, defined from SUV_{max} , was used since the main purpose was to investigate test–retest reproducibility and if the number of target lesions varies across patients, the correlation could be inflated.²¹

Metabolic activity (corrected for injected dose, decay and patient weight) was measured as SUV_{max} , peak standardized uptake value (SUV_{peak}), SUV_{mean} , the volume encompassed by the 50% SUV_{max} isocontour ($\text{MTV}_{50\%}$) and TLG from this volume. The cumulative SUV–volume histograms (CSH)³² were plotted to test reproducibility of the distribution of ^{18}F -FDG uptake

(heterogeneity) in the tumour by calculating the area under the curve (AUC).³³ The SUV and volume metrics were extracted from the software package Mirada XD3 (MIRADA Medical, Oxford, UK).

Statistical analysis

Statistical analysis was performed in SPSS® v. 19 (SPSS Inc., Chicago, IL) unless otherwise noted. Reproducibility of each measure was assessed from the difference between the two scans (Scan 1 – Scan 2), and coefficients of variation (CVs) as $\text{CVs} = \sigma/\mu$ (σ , spread; μ , mean) were assessed. For further analysis, visual assessment of P-P plots and Kolmogorov–Smirnov and Shapiro–Wilk tests were used to assess the distribution of SUV and volume measurements, and a log transformation was used where deemed necessary to match a normal distribution. Bland–Altman plots with limits of agreement from Bland–Altman analysis³⁴ were generated. Variance component analysis was performed using the VARCOMP procedure in SPSS to compare variance between measurements to variance between patients (corresponding to signal-to-noise comparison). Wilcoxon signed rank test was used to test for difference in CVs between SUV_{max} and SUV_{peak} in PET/CT (PSF/TOF) and to test the difference in CVs of SUV_{max} from PET/CT (OSEM) and PET/MR.

RESULTS

Table 1 describes the number of scans that the 30 patients received. Figure 1 shows a flowchart of inclusion. Three did not receive the second scan; one patient was ill with flu, one dropped out and one patient forgot to come for the second scans. The remaining 27 had carcinoma in the oropharynx, but 3 patients had no pathological ^{18}F -FDG uptake and were excluded from further analysis; Table 1 shows further characteristics of the remaining 24 patients. 21 of the 24 patients were also scanned twice with PET/MR, however, the data from 1 scan could not be reconstructed leaving 20 patients with 2 baseline PET/MR. All patients were scanned according to the same protocol, and the mean time difference between ^{18}F -FDG injection and scanning (uptake time Scan 1 – uptake time Scan 2) was -1.7 min (range, -10 to 4 min) for PET/CT and -5.6 min (range, -49 to 8 min) minutes for PET/MR. The difference in uptake time was very small, but on average, the uptake time was a little longer on the second day because an unplanned delay on Day 1 could be matched by the same delay on Day 2, whereas it was not possible to compensate if an unplanned delay happened on the second day.

Table 2, shows the result for lesion counting, and Table 3 shows mean values for ^{18}F -FDG measurements and volumes for both

Table 2. Lesions encountered in the 21 patients with both 2 positron emission tomography (PET)/CT scans and 2 PET/MR scans

Number of lesions	(1) Scan		(2) Scan	
	PET/CT (PSF)	PET/MR	PET/CT (PSF)	PET/MR
Lesions	51	40	53	39

PSF, point spread function.

Fewer lymph nodes were noted in PET/MR owing to a smaller field of view.

Table 3. Mean values of fluorine-18 fludeoxyglucose (^{18}F -FDG) and volume measurements for positron emission tomography (PET)/CT and PET/MR

^{18}F -FDG uptake measurements	Mean PET/CT (PSF/TOF)	Mean PET/CT (OSEM)	Mean PET/MR
SUV_{max}	20.4	15.0	16.1
SUV_{peak}	14.0	12.1	13.3
SUV_{mean}	13.3	10.2	11.0
$\text{MTV}_{50\%}$	5.0	7.1	6.6
Total lesion glycolysis	74.8	80.9	81.5

$\text{MTV}_{50\%}$, metabolic tumour volume from threshold of $\geq 50\%$ of SUV_{max} ; OSEM, ordered subset expectation maximization; PSF, point spread function; SUV_{max} , maximum standardized uptake value; SUV_{mean} , mean standardized uptake value; SUV_{peak} , peak standardized uptake value; TOF, time of flight.

PET/CT and PET/MR. Mean difference in injected ^{18}F -FDG activity between the scans was 0.01 MBq kg^{-1} .

$\text{MTV}_{50\%}$ and TLG needed to be log transformed to follow normal distribution as assessed visually and from statistical tests

($p < 0.03$) using Kolmogorov–Smirnov and Shapiro–Wilk tests. After log transformation, there was no deviation from normality. We observed an increase in variation of absolute SUV difference with increasing SUV of the two pre-treatment PET/CT scans. Consequently, the difference in Bland–Altman analysis is

Figure 2. Bland–Altman plots of the maximum standardized uptake value (SUV_{max}) and peak standardized uptake value (SUV_{peak}) for positron emission tomography (PET)/CT reconstructed with point spread function (PSF)/time of flight (TOF) (a, d) and ordered subset expectation maximization (OSEM) (b, e) and for PET/MR (c, f). The difference in percent is plotted against the average; the solid line is reference for no difference, and the dashed lines represent the mean and the upper and lower limits of agreement

Figure 3. Bland–Altman plots of metabolic tumour volume from threshold of maximum standardize uptake value ($\text{MTV}_{50\%}$) and total lesion glycolysis (TLG) for positron emission tomography (PET)/CT reconstructed with point spread function (PSF)/time of flight (TOF) (a, d) and ordered subset expectation maximization (OSEM) (b, e) and for PET/MR (c, f). $\text{MTV}_{50\%}$ and TLG were log (Ln) transformed, the y-axis is log transformed to ease the assessment of limits of agreement. The solid line is reference for no difference, and the dashed lines represent the mean and the upper and lower limits of agreement.

reported in percent of SUV, which did not appear to depend on the pre-treatment SUV. Likewise, the Bland–Altman plots were generated with difference in percent to emphasize potential difference in variance, *i.e.* increasing variance with increasing SUVs. The Bland–Altman plots for SUV_{max} and SUV_{peak} for PET/CT (PSF/TOF, OSEM) and for PET/MR can be seen from Figure 2, and Bland–Altman plots for TLG and $\text{MTV}_{50\%}$ can be seen from Figure 3. Results from Bland–Altman analysis are presented in Table 4, and Table 5 depicts CVs and the results from the variance component analysis from PET/CT (PSF/TOF, OSEM) and PET/MR. SUV_{max} , SUV_{mean} and SUV_{peak} measures were generally very reproducible with limits of agreement within approximately $\pm 20\%$ for all measures. MTV and TLG measures were also very reproducible with limits of agreement within approximately $\pm 25\%$.

Figure 4 shows an example of a PET/CT (PSF) and a PET/MR scan from the patient with the largest difference in MTV

($\text{MTV}_{50\%}$) on PET/CT. Figure 5 shows an example of a patient with same tumour size, but with a larger difference in $\text{MTV}_{50\%}$ on PET/MR than on PET/CT. The differences of 35% (Figure 4a, b) and of 13% (Figure 5c, d) are not easily distinguished from the two figures since the blue contour ($\text{MTV}_{50\%}$) obviously marks the same area in both PET/CT and PET/MR on both scan days.

Overall, the distribution of ^{18}F -FDG uptake was very reproducible for both PET/CT and PET/MR reported as AUC from CSH, and the mean difference in AUC can be seen from Table 4. The small percentage variability for the AUC measures (Table 5) was not reflected in better variance components; 74% and 82% of the total variation in the data set for PET/CT was explained by interpatient variance. For all other measures, >95% of the total variance was owing to interpatient variance (Table 4), which illustrates small CVs for all SUV measurements (approximately 5–7%) and for MTV and TLG for both PET/CT and PET/MR.

Table 4. Results from Bland–Altman analysis

¹⁸ F-FDG uptake measurements	PET/CT (PSF/TOF)			PET/CT (OSEM)			PET/MR		
	Mean difference	SD	Limits of agreement (%)	Mean difference	SD	Limits of agreement (%)	Mean difference	SD	Limits of agreement (%)
SUV _{max}	−3.7	10.3	(−24.4, 16.9)	−2.8	6.3	(−15.5, 9.8)	3.2	8.6	(−14.0, 20.4)
SUV _{peak}	−4.1	7.3	(−18.7, 10.6)	−4.5	6.9	(−18.2, 9.2)	2.2	8.0	(−13.9, 18.2)
SUV _{mean50%}	−3.5	9.5	(−22.4, 15.5)	−4.2	7.1	(−18.4, 10.0)	3.1	7.8	(−12.5, 18.7)
MTV _{50%}	N/A	N/A	(−29.7, 39.5)	N/A	N/A	(−30.9, 48.5)	N/A	N/A	(−27.7, 28.2)
TLG _{50%}	N/A	N/A	(−29.6, 29.7)	N/A	N/A	(−30.5, 36.3)	N/A	N/A	(−24.2, 30.1)
AUC unnormalized	−3.5	9.6	(−22.7, 15.6)	−4.2	7.1	(−18.4, 10.0)	3.1	7.9	(−12.7, 18.9)
AUC normalized	0.2	2.7	(−5.1, 5.5)	−1.4	2.5	(−6.3, 3.6)	−0.1	1.5	(−3.2, 2.9)

AUC, area under the curve; MTV_{50%}, metabolic tumour volume from threshold of $\geq 50\%$ of SUV_{max}; OSEM, ordered subset expectation maximization; PET, positron emission tomography; PSF, point spread function; SD, standard deviation; SUV_{max}, maximum standard uptake value; SUV_{peak}, peak standard uptake value; TLG_{50%}, total lesion glycolysis calculated as the product from MTV_{50%} and SUV_{mean} from MTV_{50%}; TOF, time of flight. The results are presented as percentages, except for MTV and TLG, where a log transformation was applied to fulfil the normal distribution assumption. The mean difference (SD) for lnMTV for PET/CT (PSF/TOF) was −0.01 (0.17), for PET/CT (OSEM) was 0.013 (0.19) and for PET/MR was −0.04 (0.14). The mean difference (SD) for lnTLG for PET/CT (PSF/TOF) was −0.01 to 0.05 (0.15), for PET/CT (OSEM) was −0.03 (0.17) and for PET/MR was −0.01 (0.13). Limits of agreement were assessed from mean difference $\pm$ 2 SD. For MTV and TLG, the limits of agreement were transformed back using the exponential function.

There was no significant difference in CVs of SUV_{max} and SUV_{peak} in PET/CT (PSF/TOF) ($p = 0.09$); neither was there any significant difference in CVs of SUV_{max} from PET/CT (OSEM) and from PET/MR ($p = 0.31$).

DISCUSSION

This is, to our knowledge, the first study of reproducibility of ¹⁸F-FDG uptake in ¹⁸F-FDG PET/CT in patients with HNSCC and the very first to test reproducibility of ¹⁸F-FDG uptake values on PET/MR. There were no distinct difference, in

lesions countered by the two PET/CT scans nor by the two PET/MR scans, except in one patient where lymph nodes were counted on the second scan, but not on the first scan, and this was only in PET/CT. In the remaining 20 patients, difference in lesions was because lymph nodes conglomerate was counted as two lesions in one scan and one lesion in the other scan. We observed no difference in primary tumours and no difference in lymph nodes side of the neck, *i.e.* ipsilateral and contralateral. The difference between lesions found in PET/CT and PET/MR were owing to the smaller field of view

Table 5. Coefficients of variation (CVs) in percent for positron emission tomography (PET)/CT reconstructed with point spread function (PSF) and with ordered subset expectation maximization (OSEM) and for PET/MR

¹⁸ F-FDG uptake measurements	PET/CT (PSF/TOF) CV (%)	Variance component within subjects (%)	PET/CT (OSEM) CV (%)	Variance component within subjects (%)	PET/MR CV (%)	Variance component within subjects (%)
SUV _{max}	7.6	3.4	4.8	1.5	6.4	3.6
SUV _{peak}	5.8	1.7	5.7	1.6	5.7	2.6
SUV _{mean50%}	7.0	2.6	5.7	1.6	5.8	2.7
MTV _{50%} ^a	11.8	1.8	11.2	2.7	9.2	2.5
TLG _{50%} ^a	11.0	1.0	11.7	1.6	10.2	1.0
AUC unnormalized	7.1	2.6	5.75	1.6	5.87	2.7
AUC normalized	1.9	25.9	2.0	17.9	1.1	4.2

AUC, area under the curve; MTV_{50%}, metabolic tumour volume from threshold of $\geq 50\%$ of SUV_{max}; SUV_{max}, maximum standard uptake value; SUV_{peak}, peak standard uptake value; TLG_{50%}, total lesion glycolysis calculated as the product from MTV_{50%} and SUV_{mean} from MTV_{50%}; TOF, time of flight. Variance of components within subjects was assessed from a variance component analysis and defined as total variance – variance between subjects.

^aVariance component was carried out after log transformation.

Figure 4. An example each from the positron emission tomography (PET)/CT (point spread function) and PET/MR scans of a patient with 35% difference in metabolic tumour volume ($\text{MTV}_{50\%}$) (from 9.5 to 6.1 cm^3) and 27% in maximum standardized uptake value (SUV_{max}) (from 14.7 to 11.8) in PET/CT. In PET/MR, the difference in $\text{MTV}_{50\%}$ was 3% (from 11.5 to 11.2 cm^3) and 2% in SUV_{max} (from 12.1 to 12.4). (a, b) A PET/CT slice from the first and second scans, respectively, and (c, d) a PET/MR slice each from the first and second scans, respectively.

in PET/MR. ^{18}F -FDG uptake measured as SUVs (SUV_{max} , SUV_{peak} , SUV_{mean}) or as AUC from cumulative CSH is demonstrated to be very reproducible for both PET/CT and PET/MR.

Functional imaging with PET/CT, and in the future perhaps also with PET/MR, is increasingly used in the management of HNSCC for staging, radiotherapy planning and/or treatment response evaluation. ^{18}F -FDG PET/CT is useful in patients with HNSCC for prognostication, target delineation, response assessment and possibly for dose painting. However, for appropriate implementation of all four applications, the reproducibility of ^{18}F -FDG uptake must be known. In particular, if ^{18}F -FDG PET images are used for dose painting with dose modification according to ^{18}F -FDG uptake, the distribution of ^{18}F -FDG uptake must be stable and reproducible before the start of radiotherapy.

Our study shows reasonably good reproducibility (*i.e.*, a scan-to-scan variation of 20%) of the ^{18}F -FDG uptake measurements on both PET/CT (PSF) and PET/MR. Of note, the present study

represents the ideal scenario where the same scanners and protocols are used in the comparisons. If different scanners are used, additional margins for uncertainty may have to be included.³⁵ Figure 4 depicts the worst-case scenario in difference in $\text{MTV}_{50\%}$ in PET/CT (PSF), and such difference is not easy to see from the figure and the large relative difference is also owing to the actual tumour size, where a difference in 1 cm^3 would yield a relative difference of 10%. Furthermore, the blue contour obviously marks the same area in both scans on both PET/CT (PSF) and PET/MR.

Most of the SUV metrics do not assess a change in tracer uptake within the entire tumour, and there is increasing awareness of intratumoral heterogeneity in ^{18}F -FDG uptake.³⁶ The distribution of ^{18}F -FDG uptake assessed by CSH (with or without normalization to SUV_{max}) is very reproducible for both PET/CT and PET/MR. However, although the small CV for the normalized AUC seems very appealing, it should be noted from Table 5 that there is a relatively small interpatient variance component. This is because the good reproducibility of the

Figure 5. An example each from the positron emission tomography (PET)/CT (point spread function) and PET/MR scans of a patient with 1% difference in metabolic tumour volume ($MTV_{50\%}$) in PET/CT (from 9.5 to 9.6 cm^3) and 11% in maximum standardized uptake value (SUV_{max}) (from 30.0 to 33.6). In PET/MR, the difference in $MTV_{50\%}$ was 13% (from 12.1 to 13.7 cm^3) and 4% in SUV_{max} (from 27.7 to 26.7). (a, b) A PET/CT slice from the first and second scans, respectively, and (c, d) a PET/MR slice each from the first and second scans, respectively.

normalized AUC (low noise) is outweighed by a narrow range of variation between patients (low signal).

The variability in the present reproducibility study is within the test–retest variability reported in the meta-analysis by de Langen *et al.*⁸ They reported that a difference should exceed both absolute and relative thresholds of 3.1% and 25% for SUV_{max} and 1.2% and 20% for SUV_{mean} to be called a “true difference”. According to the proposed PERCIST response criteria,⁹ a change in $\pm 30\%$ in SUL_{peak} (lean body mass corrected SUV) and an absolute change, at least 0.8 SUL, is classified as progression or remission. The worst-case scenario patient from Figure 4 had an increase in SUV_{max} from 14.7 to 19.2 (a change of 4.5 corresponding to 27%), however, the change in SUV_{peak} and SUV_{mean} was $<20\%$. Of note, the increase in SUV_{max} actually yields a decrease in $MTV_{50\%}$, illustrating that it is challenging to separate

a metabolic response assessment from a structural response assessment, however, this discussion is beyond the scope of this article.

All patients in the present test–retest study would have been assessed as stable disease using the PERCIST criteria, although metabolic activity was assessed as SUV and not SUL. They would be expected to be similar since it is unlikely that a patient would change composition of body fat and lean tissue without a change in weight within 3 days. It must of course be noted that our patient cohort is very homogenous and only includes patients with HNSCC, and it is possible that a larger variation would be encountered in clinical practice with different tumour types. PERCIST proposes that no more than five lesions with a minimal ^{18}F -FDG uptake of approximately 2.5⁹ (calculated as $1.5 \times \text{the liver } SUV_{mean} + 2 \text{ standard deviations}$) be evaluated. In contrast to PERCIST, only one lesion

per patient was assessed in the present test–retest study to avoid introducing bias in the evaluation of reproducibility if the number of lesions varied from patient to patient.

We observed an increase in absolute value variation with increasing SUVs. Consequently, reporting or using absolute threshold values for SUVs would result in too narrow limits at low SUV and too wide limits at high SUV, and this is not recommended in PERCIST criteria⁹ nor in the meta-analysis mentioned above.

Our scans were obtained 3 days apart, and therefore the fundamental assumption behind the study was that there should be no appreciable difference in tumour characteristics between the two scans, *i.e.* the true result is “stable disease”. Reassuringly, the values from PERCIST and the de Langen study would indeed classify all our patients as stable disease.

CONCLUSION

The ^{18}F -FDG uptake in ^{18}F -FDG PET/CT scans and ^{18}F -FDG PET/MR scans are both very reproducible in patients with HNSCC. The uptake in PET/MR scans is as reproducible as in PET/CT scans. The analyses did not indicate that SUV_{peak} is more reproducible than SUV_{max} in neither PET/CT nor PET/MR. Pre-therapy ^{18}F -FDG uptake is sufficiently stable in individual tumours to allow treatment modification on this basis.

FUNDING

The corresponding author has received a grant from the Arvid Nilssons Foundation. Supported by the Global Excellence in Health program of the Capital Region of Denmark.

ACKNOWLEDGMENTS

The authors thank Søren M Bentzen for suggesting the variance component analysis.

REFERENCES

- Xie P, Li M, Zhao H, Sun X, Fu Z, Yu J. ^{18}F -FDG PET or PET-CT to evaluate prognosis for head and neck cancer: a meta-analysis. *J Cancer Res Clin Oncol* 2011; **137**: 1085–93. doi: [10.1007/s00432-010-0972-y](https://doi.org/10.1007/s00432-010-0972-y)
- Zhang B, Li X, Lu X. Standardized uptake value is of prognostic value for outcome in head and neck squamous cell carcinoma. *Acta Otolaryngol* 2010; **130**: 756–62. doi: [10.3109/00016480903402981](https://doi.org/10.3109/00016480903402981)
- Rasmussen JH, Vogelius IR, Fischer BM, Friberg J, Aznar MC, Persson GF, et al. Prognostic value of ^{18}F -fluorodeoxyglucose uptake in 287 patients with head and neck squamous cell carcinoma. *Head Neck* May 2014. Epub ahead of print. doi: [10.1002/hed.23745](https://doi.org/10.1002/hed.23745)
- Troost EG, Schinagel DA, Bussink J, Oyen WJ, Kaanders JH. Clinical evidence on PET-CT for radiation therapy planning in head and neck tumours. *Radiother Oncol* 2010; **96**: 328–34. doi: [10.1016/j.radonc.2010.07.017](https://doi.org/10.1016/j.radonc.2010.07.017)
- Troost EG, Schinagel DA, Bussink J, Boerman OC, van der Kogel AJ, Oyen WJ, et al. Innovations in radiotherapy planning of head and neck cancers: role of PET. *J Nucl Med* 2010; **51**: 66–76. doi: [10.2967/jnumed.108.061499](https://doi.org/10.2967/jnumed.108.061499)
- Madani I, Duprez F, Boterberg T, Van de Wiele C, Bonte K, Deron P, et al. Maximum tolerated dose in a phase I trial on adaptive dose painting by numbers for head and neck cancer. *Radiother Oncol* 2011; **101**: 351–5. doi: [10.1016/j.radonc.2011.06.020](https://doi.org/10.1016/j.radonc.2011.06.020)
- Vogelius IR, Håkansson K, Due AK, Aznar MC, Berthelsen AK, Kristensen CA, et al. Failure-probability driven dose painting. *Med Phys* 2013; **40**: 081717. doi: [10.1118/1.4816308](https://doi.org/10.1118/1.4816308)
- de Langen AJ, Vincent A, Velasquez LM, van Tinteren H, Boellaard R, Shankar LK, et al. Repeatability of ^{18}F -FDG uptake measurements in tumors: a metaanalysis. *J Nucl Med* 2012; **53**: 701–8. doi: [10.2967/jnumed.111.095299](https://doi.org/10.2967/jnumed.111.095299)
- 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**(Suppl. 1): 122S–50S. doi: [10.2967/jnumed.108.057307](https://doi.org/10.2967/jnumed.108.057307)
- Choi ES, Ha SG, Kim HS, Ha JH, Paeng JC, Han I. Total lesion glycolysis by ^{18}F -FDG PET/CT is a reliable predictor of prognosis in soft-tissue sarcoma. *Eur J Nucl Med Mol Imaging* 2013; **40**: 1836–42. doi: [10.1007/s00259-013-2511-y](https://doi.org/10.1007/s00259-013-2511-y)
- Moon SH, Choi JY, Lee HJ, Son YI, Baek CH, Ahn YC, et al. Prognostic value of ^{18}F -FDG PET/CT in patients with squamous cell carcinoma of the tonsil: comparisons of volume-based metabolic parameters. *Head Neck* 2013; **35**: 15–22. doi: [10.1002/hed.22904](https://doi.org/10.1002/hed.22904)
- Hyun SH, Ahn HK, Kim H, Ahn MJ, Park K, Ahn YC, et al. Volume-based assessment by (^{18}F)-FDG PET/CT predicts survival in patients with stage III non-small-cell lung cancer. *Eur J Nucl Med Mol Imaging* 2014; **41**: 50–8. doi: [10.1007/s00259-013-2530-8](https://doi.org/10.1007/s00259-013-2530-8)
- Lee JW, Cho A, Lee JH, Yun M, Lee JD, Kim YT, et al. The role of metabolic tumor volume and total lesion glycolysis on ^{18}F -FDG PET/CT in the prognosis of epithelial ovarian cancer. *Eur J Nucl Med Mol Imaging* 2014; **41**: 1898–906. doi: [10.1007/s00259-014-2803-x](https://doi.org/10.1007/s00259-014-2803-x)
- Fletcher JW, Djulbegovic B, Soares HP, Siegel BA, Lowe VJ, Lyman GH, et al. Recommendations on the use of ^{18}F -FDG PET in oncology. *J Nucl Med* 2008; **49**: 480–508. doi: [10.2967/jnumed.107.047787](https://doi.org/10.2967/jnumed.107.047787)
- Hatt M, Cheze-Le Rest C, Aboagye EO, Kenny LM, Rosso L, Turkheimer FE, et al. Reproducibility of ^{18}F -FDG and 3'-deoxy-3'- ^{18}F -fluorothymidine PET tumor volume measurements. *J Nucl Med* 2010; **51**: 1368–76. doi: [10.2967/jnumed.110.078501](https://doi.org/10.2967/jnumed.110.078501)
- Lindholm H, Staaf J, Jacobsson H, Brolin F, Hatherly R, Sánchez-Crespo A. Repeatability of the maximum standard uptake value (SUV_{max}) in FDG PET. *Mol Imaging Radionucl Ther* 2014; **23**: 16–20. doi: [10.4274/Mirt.76376](https://doi.org/10.4274/Mirt.76376)
- Nahmias C, Wahl LM. Reproducibility of standardized uptake value measurements determined by ^{18}F -FDG PET in malignant tumors. *J Nucl Med* 2008; **49**: 1804–8. doi: [10.2967/jnumed.108.054239](https://doi.org/10.2967/jnumed.108.054239)
- Tixier F, Hatt M, Le Rest CC, Le Pogam A, Corcos L, Visvikis D. Reproducibility of tumor uptake heterogeneity characterization through textural feature analysis in ^{18}F -FDG PET. *J Nucl Med* 2012; **53**: 693–700. doi: [10.2967/jnumed.111.099127](https://doi.org/10.2967/jnumed.111.099127)
- Velasquez LM, Boellaard R, Kolia G, Hayes W, Hoekstra OS, Lammertsma AA, et al. Repeatability of ^{18}F -FDG PET in a multicenter phase I study of patients with advanced gastrointestinal malignancies. *J Nucl Med* 2009; **50**: 1646–54. doi: [10.2967/jnumed.109.063347](https://doi.org/10.2967/jnumed.109.063347)
- Weber WA, Ziegler SI, Thödtmann R, Hanauske AR, Schwaiger M. Reproducibility of metabolic measurements in malignant

- tumors using FDG PET. *J Nucl Med* 1999; **40**: 1771–7.
21. Rockall AG, Avril N, Lam R, Iannone R, Mozley PD, Parkinson C, et al. 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. doi: [10.1158/1078-0432.CCR-13-2634](https://doi.org/10.1158/1078-0432.CCR-13-2634)
 22. Becker M, Zaidi H. Imaging in head and neck squamous cell carcinoma: the potential role of PET/MRI. *Br J Radiol* 2014; **87**: 20130677. doi: [10.1259/bjr.20130677](https://doi.org/10.1259/bjr.20130677)
 23. Varoquaux A, Rager O, Poncet A, Delattre BM, Ratib O, Becker CD, et al. Detection and quantification of focal uptake in head and neck tumours: (18)F-FDG PET/MR versus PET/CT. *Eur J Nucl Med Mol Imaging* 2014; **41**: 462–75. doi: [10.1007/s00259-013-2580-y](https://doi.org/10.1007/s00259-013-2580-y)
 24. Queiroz MA, Hüllner M, Kuhn F, Huber G, Meerwein C, Kollias S, et al. PET/MRI and PET/CT in follow-up of head and neck cancer patients. *Eur J Nucl Med Mol Imaging* 2014; **41**: 1066–75. doi: [10.1007/s00259-014-2707-9](https://doi.org/10.1007/s00259-014-2707-9)
 25. Platzek I, Beuthien-Baumann B, Schneider M, Gudziol V, Langner J, Schramm G, et al. PET/MRI in head and neck cancer: initial experience. *Eur J Nucl Med Mol Imaging* 2013; **40**: 6–11. doi: [10.1007/s00259-012-2248-z](https://doi.org/10.1007/s00259-012-2248-z)
 26. Kubiessa K, Purz S, Gawlitza M, Kühn A, Fuchs J, Steinhoff KG, et al. Initial clinical results of simultaneous 18F-FDG PET/MRI in comparison to 18F-FDG PET/CT in patients with head and neck cancer. *Eur J Nucl Med Mol Imaging* 2014; **41**: 639–48. doi: [10.1007/s00259-013-2633-2](https://doi.org/10.1007/s00259-013-2633-2)
 27. Buchbender C, Heusner TA, Lauenstein TC, Bockisch A, Antoch G. Oncologic PET/MRI, part 1: tumors of the brain, head and neck, chest, abdomen, and pelvis. *J Nucl Med* 2012; **53**: 928–38. doi: [10.2967/jnumed.112.105338](https://doi.org/10.2967/jnumed.112.105338)
 28. Boss A, Stegger L, Bisdas S, Kolb A, Schwenzer N, Pfister M, et al. Feasibility of simultaneous PET/MR imaging in the head and upper neck area. *Eur Radiol* 2011; **21**: 1439–46. doi: [10.1007/s00330-011-2072-z](https://doi.org/10.1007/s00330-011-2072-z)
 29. Jakoby BW, Bercier Y, Conti M, Casey ME, Bendriem B, Townsend DW. Physical and clinical performance of the mCT time-of-flight PET/CT scanner. *Phys Med Biol* 2011; **56**: 2375–89. doi: [10.1088/0031-9155/56/8/004](https://doi.org/10.1088/0031-9155/56/8/004)
 30. Panin VY, Kehren F, Michel C, Casey M. Fully 3-D PET reconstruction with system matrix derived from point source measurements. *IEEE Trans Med Imaging* 2006; **25**: 907–21.
 31. Lois C, Jakoby BW, Long MJ, Hubner KF, Barker DW, Casey ME, et al. An assessment of the impact of incorporating time-of-flight information into clinical PET/CT imaging. *J Nucl Med* 2010; **51**: 237–45. doi: [10.2967/jnumed.109.068098](https://doi.org/10.2967/jnumed.109.068098)
 32. El Naqa I, Grigsby P, Apte A, Kidd E, Donnelly E, Khullar D, et al. Exploring feature-based approaches in PET images for predicting cancer treatment outcomes. *Pattern Recognit* 2009; **42**: 1162–71.
 33. van Velden FH, Cheebsumon P, Yaqub M, Smit EF, Hoekstra OS, Lammertsma AA, et al. Evaluation of a cumulative SUV-volume histogram method for parameterizing heterogeneous intratumoural FDG uptake in non-small cell lung cancer PET studies. *Eur J Nucl Med Mol Imaging* 2011; **38**: 1636–47. doi: [10.1007/s00259-011-1845-6](https://doi.org/10.1007/s00259-011-1845-6)
 34. Sedgwick P. Limits of agreement (Bland-Altman method). *BMJ* 2013; **346**: f1630. doi: [10.1136/bmj.f1630](https://doi.org/10.1136/bmj.f1630)
 35. Kumar V, Nath K, Berman CG, Kim J, Tanvetyanon T, Chiappori AA, et al. Variance of SUVs for FDG-PET/CT is greater in clinical practice than under ideal study settings. *Clin Nucl Med* 2013; **38**: 175–82. doi: [10.1097/RLU.0b013e318279ffdf](https://doi.org/10.1097/RLU.0b013e318279ffdf)
 36. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA Jr, Kinzler KW. Cancer genome landscapes. *Science* 2013; **339**: 1546–58. doi: [10.1126/science.1235122](https://doi.org/10.1126/science.1235122)