


Repeatability of tumour hypoxia imaging using [¹⁸F]EF5 PET/CT in head and neck cancer

Antti Silvonemi^{1,2}  · Sami Suilamo^{3,4} · Timo Laitinen¹ · Sarita Forsback¹ ·
Eliisa Löyttyniemi⁵ · Samuli Vaittinen⁶ · Virva Saunavaara^{1,4} · Olof Solin¹ ·
Tove J. Grönroos^{1,3} · Heikki Minn^{1,3}

Received: 8 September 2017 / Accepted: 9 October 2017 / Published online: 26 October 2017
© The Author(s) 2017. This article is an open access publication

Abstract

Purpose Hypoxia contributes to radiotherapy resistance and more aggressive behaviour of several types of cancer. This study was designed to evaluate the repeatability of intratumour uptake of the hypoxia tracer [¹⁸F]EF5 in paired PET/CT scans.

Methods Ten patients with newly diagnosed head and neck cancer (HNC) received three static PET/CT scans before chemoradiotherapy: two with [¹⁸F]EF5 a median of 7 days apart and one with [¹⁸F]FDG. Metabolically active primary tumour volumes were defined in [¹⁸F]FDG images and transferred to co-registered [¹⁸F]EF5 images for repeatability analysis. A tumour-to-muscle uptake ratio (TMR) of 1.5 at 3 h from injection of [¹⁸F]EF5 was used as a threshold representing hypoxic tissue.

Results In 10 paired [¹⁸F]EF5 PET/CT image sets, SUVmean, SUVmax, and TMR showed a good correlation with the intraclass correlation coefficients of 0.81, 0.85, and 0.87, respectively. The relative coefficients of repeatability for these

parameters were 15%, 17%, and 10%, respectively. Fractional hypoxic volumes of the tumours in the repeated scans had a high correlation using the Spearman rank correlation test ($r = 0.94$). In a voxel-by-voxel TMR analysis between the repeated scans, the mean of Pearson correlation coefficients of individual patients was 0.65. The mean (\pm SD) difference of TMR in the pooled data set was 0.03 ± 0.20 .

Conclusion Pretreatment [¹⁸F]EF5 PET/CT within one week shows high repeatability and is feasible for the guiding of hypoxia-targeted treatment interventions in HNC.

Keywords Repeatability · Hypoxia · Pet · Head and neck cancer · ¹⁸F-EF5

Introduction

Hypoxia is among the strongest biological factors causing radiotherapy (RT) resistance in several types of cancer [1]. The outcome of RT in head and neck cancer (HNC) is particularly sensitive to the oxygenation status where numerous studies show a worse prognosis associated with tumour hypoxia [2, 3]. To overcome radiation resistance, dose-escalation protocols have been developed where hypoxic subvolumes receive higher doses based on PET/CT imaging with tracers preferentially accumulating in poorly oxygenated tissues [4]. One of these PET tracers is a fluorine-18 labelled form of 2-(2-nitro-1H-imidazol-1-yl)-N-(2,2,3,3,3-pentafluoropropyl)-acetamide (EF5), which has been thoroughly evaluated for its hypoxia-avidity potential both in vitro and in vivo [5, 6]. Our previous studies have shown favourable tumour uptake characteristics [7] and a prognostic value of [¹⁸F]EF5 PET/CT imaging in patients with HNC [8].

The repeatability of hypoxia PET imaging is crucial for feasibility in clinical applications such as the planning of RT.

✉ Antti Silvonemi
anmisi@utu.fi

¹ Turku PET Centre, University of Turku, P.O. BOX 52, FI-20521 Turku, Finland

² Department of Otorhinolaryngology – Head and Neck Surgery, Turku University Hospital, P.O. BOX 52, FI-20521 Turku, Finland

³ Department of Oncology and Radiotherapy, Turku University Hospital, P.O. BOX 52, FI-20521 Turku, Finland

⁴ Department of Medical Physics, Turku University Hospital, P.O. BOX 52, FI-20521 Turku, Finland

⁵ Department of Biostatistics, University of Turku, Turku, Finland

⁶ Department of Pathology, Turku University Hospital, P.O. BOX 52, FI-20521 Turku, Finland

A few clinical studies have been conducted previously using voxel-by-voxel analysis of spatial tracer distribution in HNC and lung cancer in paired scans within a short pretreatment period with [^{18}F]FMISO PET/CT [9–11] and [^{18}F]HX4 PET/CT [12]. Three of these studies reported highly repeatable results of hypoxia PET imaging [10–12], while a single study showed a lower linear correlation than the other studies between the intratumour tracer uptake in the repeated scans [9]. These studies were performed using a fixed interval of two [10] or three [9] days or an average interval of 1–2 days [11, 12], respectively.

Owing to the potential differences in biodistribution and imaging characteristics of the available hypoxia-avid PET tracers, it is important to investigate their repeatability individually in the test-retest setting. Therefore, we aimed to measure the repeatability of [^{18}F]EF5 PET/CT imaging in HNC before the start of RT and focus on the possibility of comparing our findings to [^{18}F]FMISO and [^{18}F]HX4, which have been evaluated in comparable studies and patients.

Patients and methods

Patients

This prospective study (NCT 01774760) was conducted at Turku University Hospital, Finland between September 2013 and September 2016. Patients between 18 and 80 years of age with untreated pharyngeal squamous cell carcinoma referred to definitive chemoradiotherapy (CRT) were eligible and were required to have a WHO performance status 0–2, without a history of previous head and neck malignancies or RT in the head and neck area. Additional exclusion criteria were serious cardiac, pulmonary, renal, liver and haematological disorders, pregnancy, or nursing. Eleven patients signed the consent form and underwent all study procedures, but one tonsillar cancer patient was excluded from analyses due to an extremely low tracer uptake in the primary tumour in pretreatment [^{18}F]FDG PET/CT imaging after diagnostic tonsillectomy, which preceded PET/CT imaging. The characteristics of the remaining 10 patients are presented in Table 1. Four of them had primary tumour positive for p16 which has been linked to human papilloma virus (HPV) infection [13].

Synthesis of [^{18}F]EF5

[^{18}F]EF5 was synthesised as previously described [14]. The molar activity of [^{18}F]EF5 decay corrected to the end of synthesis exceeded 8 GBq/ μmol . Radiochemical purity was higher than 98.5% in every production batch.

Imaging protocol

All patients underwent two [^{18}F]EF5 PET/CT acquisitions with a median interval of 7 days (range 5–7 days). The first (EF5#1) and the second (EF5#2) scan were performed with identical acquisition protocols using the same scanner GE D690 PET/CT (General Electric Medical Systems, Milwaukee, WI, USA). Before each of the paired acquisitions, the patients received an intravenous mean (\pm SD) dose of [^{18}F]EF5 of 303 ± 23 MBq (range 246–345 MBq). The mean (\pm SD) inpatient difference between injected doses was 18 ± 12 MBq (range 1–40 MBq). A low-dose CT (120 kV, noise index 20, Asir 40%) for anatomical reference and attenuation correction was obtained immediately before PET acquisition, which started 178 ± 9 min post injection (range 160–190 min). The PET acquisition time was 6 min covering an axial field of view (FOV) of 15 cm with a slice thickness of 3.27 mm. The inpatient difference between the start of the acquisition time of repeated [^{18}F]EF5 PET/CT scans was 7 ± 6 min (range 0–19 min). The patients were immobilised on the flat scanner table using a thermoplastic mask. Venous blood samples were taken before and after the imaging session, and blood activity at the mid-point of image acquisition was calculated using linear interpolation with decay-corrected blood activity values.

On a separate day, all patients underwent whole-body ^{18}F -fluorodeoxyglucose (FDG) PET/CT imaging following the standard institutional protocol used in RT planning [15]. The sequential scans were performed in random order with [^{18}F]FDG PET/CT either between or after the two [^{18}F]EF5 PET/CTs. [^{18}F]FDG PET/CT was performed with the same GE D690 scanner as the hypoxia scans, except for patients nr 1, 2, 5, and 7, who were imaged with the Discovery VCT PET/CT scanner (General Electric Medical Systems, Milwaukee, WI, USA).

The GE D690 PET/CT scanner images were reconstructed using a 192×192 matrix with a transaxial FOV of 70 cm. In order to achieve a uniform voxel size ($3.65 \times 3.65 \times 3.27$ mm) for all PET images, a corresponding 128×128 reconstruction matrix with a transaxial FOV of 46.7 cm was selected for the GE Discovery VCT PET/CT.

Image analysis

Varian Eclipse software version 13.6 (Varian Medical Systems, Palo Alto, CA, USA) was used for the determination of tracer uptake in primary tumour and reference tissue in all PET images. The delineation of the primary tumour volume of interest (VOI) was based on the metabolically active tumour volume (MATV) in the [^{18}F]FDG image using either a threshold of 40% of SUVmax or a fixed SUV 5.0 threshold, depending on which more closely matched the CT-based anatomical gross tumour volume (GTV). Posterior neck muscles were

Table 1 Patient characteristics. All patients in this study were men

Patient nr	Age (years)	Weight (kg)	Tumour site	TNM	Stage	HPV status ^a	GTV ^b (cm ³)
1	73	77	Base of tongue	T2N1M0	III	+	8.9
2	68	49	Base of tongue	T3N0M0	III	–	40.3
3	56	86	Base of tongue	T3N1M0	III	+	51.9
4	69	66	Oropharyngeal wall	T4aN1M0	IVa	–	17.0
5	58	125	Base of tongue	T2N2cM0	IVa	–	20.5
6	60	84	Nasopharynx	T4N3bM0	IVb	+	94.0
7	69	94	Hypopharynx	T4bN2cM0	IVb	–	52.4
8	76	78	Nasopharynx	T4N2bM0	IVa	–	73.5
9	66	77	Tonsil	T4aN2cM0	IVa	+	30.4
10	66	85	Hypopharynx	T4aN2bM0	IVa	–	25.4

^a Obtained from immunohistochemical analysis for presence of cyclin dependent kinase inhibitor 2A (p16) in tumour cells

^b GTV = gross tumour volume manually delineated in CT image

used as reference tissue for tracer uptake [7]. The [¹⁸F]FDG and [¹⁸F]EF5 images were rigidly registered using anatomical information from CT images.

Carimas 2.9 software (www.turkupetcentre.fi/carimas) was used for voxel-by-voxel analysis of [¹⁸F]EF5 PET/CT images. The transformation matrices were applied to the [¹⁸F]EF5 images to define the MATV-based primary tumour VOIs in [¹⁸F]EF5 images. The accuracy of VOI structure transformations for voxel-by-voxel analyses was controlled using visual inspection of images and cross tabulation of tumour uptake values.

The uptake was measured as kBq/mL and then decay corrected and converted to standardised uptake values (SUV) under the assumption of water density. Tumour SUVmean and SUVmax uptakes were determined, as well as mean uptake in posterior neck muscle reference (SUVmuscle). Hypoxic subvolume of the tumour (HV) was determined using a tumour-to-muscle uptake ratio (TMR) of 1.5 as a threshold for hypoxia [7]. Fractional hypoxic volume (FHV) was calculated by dividing the number of hypoxic voxels with the total number of voxels within tumour VOI.

Statistical analysis

Data expressed with plus/minus indicates mean and standard deviation (SD). A paired T-test was used for comparison of injected doses, injected doses per weight, and acquisition starting times of individual patients between EF5#1 and EF5#2. Correlations of normally distributed tumour-level parameters were assessed by calculating intraclass correlation coefficients (ICC). For FHVs and HVs, a non-parametric Spearman rank correlation test was used. Pearson correlation coefficients were calculated for repeated voxel-level uptake parameters. Bland-Altman plots were constructed for agreement analysis of both the tumour- and voxel-level parameters.

In addition, upper and lower limits of agreement (LoA) and coefficient of repeatability (CoR) were calculated. $p < 0.05$ was used as a level of significance (two-tailed). For the test of normality, the Shapiro-Wilk test was used for tumour-level parameters and visual assessment for voxel-level parameters. The statistical analyses were performed using SAS software version 9.4 (SAS institute, Cary, NC, USA).

Results

Anatomical GTVs of the primary tumours measured in the CT images varied considerably. The average GTV was 41.4 ± 26.9 cm³ (range 8.9–94 cm³) and the corresponding MATV was 39.0 ± 26.7 cm³ (range 7.2–100 cm³). There were no statistically significant differences between injected doses, injected doses per weight, and scanning start times of individual patients within repeated [¹⁸F]EF5 PET/CT scans (for all comparisons $p > 0.36$). Examples of two paired [¹⁸F]EF5 PET/CT images are shown in Fig. 1.

The average whole tumour uptake of [¹⁸F]EF5 (SUVmean) among all patients was at the same level in the first scan (1.49 ± 0.16) and in the second scan (1.54 ± 0.21). The corresponding values for the highest uptake (SUVmax) were 2.12 ± 0.34 and 2.09 ± 0.35 , respectively (Table 2). The correlation of these parameters within individual patients in repeated scans was high. The ICCs were 0.81 ($p < 0.001$) for SUVmean and 0.85 ($p < 0.001$) for SUVmax. The mean differences of SUVmean and SUVmax were 0.05 ± 0.11 and -0.02 ± 0.20 , respectively. Bland-Altman plots of these parameters are presented in Fig. 2. The relative CoRs for SUVmean and SUVmax were 15% and 17%, respectively.

SUVmuscle was stable within individual patients between the repeated scans (Table 2). The ICC for SUVmuscle was

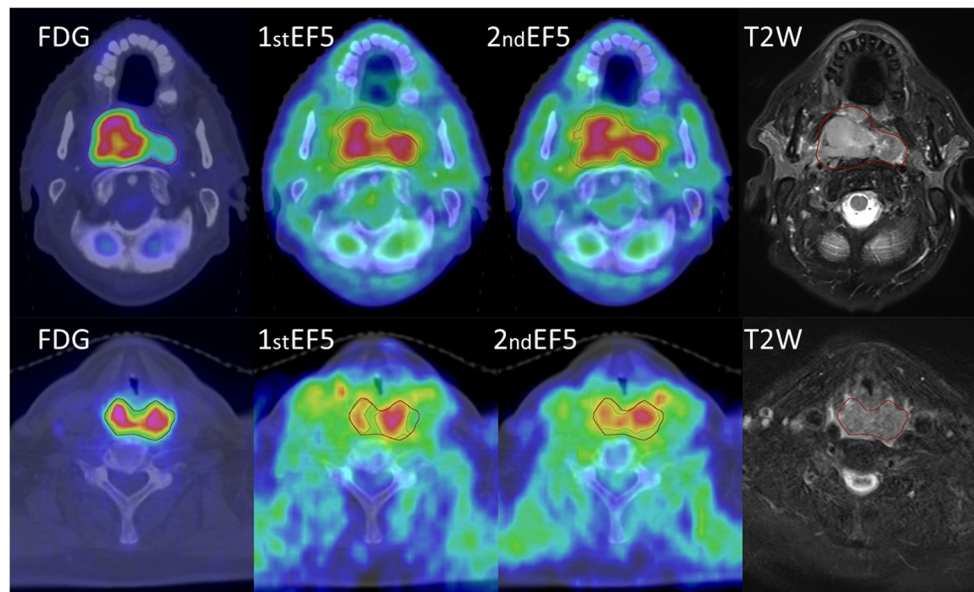


Fig. 1 PET/CT and MR images of patients presenting with nasopharyngeal cancer (No. 6; upper row) and hypopharyngeal cancer (No. 7; lower row). From left to right, corresponding axial slices from diagnostic [^{18}F]FDG, the first and the second [^{18}F]EF5 PET/CT; and fat-suppressed T2-weighted MR images are shown. The red line denotes the metabolically active tumour volume delineation using SUV 5.0 as a threshold in the [^{18}F]FDG PET image. The black line indicates hypoxic

subvolume delineation using a tumour-to-muscle uptake ratio of 1.5 as a threshold in the [^{18}F]EF5 PET image. The intrapatient voxel-by-voxel analysis showed a high correlation and agreement between the paired [^{18}F]EF5 PET/CT images for patient No. 6, while those for patient No. 7 were among the lowest of 10 patients (see results of individual patients in Fig. 3 and Tables 2 and 3)

0.84 ($p < 0.001$) and the mean difference of $\text{SUV}_{\text{muscle}}$ between the scans was 0.15 ± 0.06 with the upper and lower LoA of 0.26 and 0.04, respectively. The relative CoR for $\text{SUV}_{\text{muscle}}$ was 10%. Additionally, activity in venous blood samples measured as SUV showed a high correlation and agreement between the scans, with ICC being 0.94 ($p < 0.001$) and relative CoR 10%.

The highest correlation and agreement among tumour-level uptake parameters were observed within those of TMR. The

ICC for TMR was 0.87 ($p < 0.001$) and the mean difference was 0.02 ± 0.07 , with the upper and lower LoA of 0.17 and -0.12 , respectively (Fig. 2). The relative CoR for TMR was 10%. The parameters of HV and FHV did not show normal distribution, and therefore the correlation of these parameters between repeated scans was assessed with a Spearman correlation test. A high and statistically significant correlation was observed, both for repeated HV ($r = 0.93$, $p < 0.001$) and FHV ($r = 0.94$, $p < 0.001$).

Table 2 Tumour-level and muscle uptake parameters in the repeated [^{18}F]EF5 PET/CT scans

Patient nr	SUV_{mean} Scan 1	SUV_{mean} Scan2	SUV_{max} Scan 1	SUV_{max} Scan2	TMR Scan 1	TMR Scan 2	FHV% Scan 1	FHV% Scan 2	HV Scan 1 (cm^3)	HV Scan 2 (cm^3)	$\text{SUV}_{\text{muscle}}$ Scan 1	$\text{SUV}_{\text{muscle}}$ Scan 2
1	1.61	1.75	1.99	2.19	1.43	1.49	21.2	34.5	1.5	2.5	1.12	1.18
2	1.26	1.24	1.57	1.52	1.30	1.28	3.8	1.2	1.3	0.4	0.97	0.97
3	1.59	1.58	2.08	2.06	1.29	1.28	6.8	3.5	3.4	1.8	1.23	1.24
4	1.27	1.22	1.83	1.75	1.17	1.18	4.9	4.3	1.0	0.9	1.09	1.03
5	1.55	1.49	2.33	1.95	1.34	1.27	19.1	6.7	4.0	1.4	1.15	1.18
6	1.65	1.68	2.46	2.59	1.52	1.55	40.5	45.2	40.6	45.3	1.08	1.09
7	1.40	1.64	2.67	2.43	1.46	1.48	34.2	40.3	18.3	21.6	0.96	1.10
8	1.71	1.89	2.40	2.53	1.45	1.64	32.6	63.7	16.4	32.0	1.18	1.15
9	1.38	1.48	1.85	2.05	1.43	1.47	27.8	32.4	10.4	12.1	0.96	1.01
10	1.52	1.43	2.02	1.87	1.21	1.17	1.5	0.3	0.3	0.0	1.25	1.22
Mean	1.49	1.54	2.12	2.09	1.36	1.38	19.3	23.2	9.7	11.8	1.10	1.12
SD	0.16	0.21	0.34	0.35	0.12	0.16	14.3	22.8	12.7	16.0	0.11	0.09

TMR tumour-to-muscle uptake ratio, FHV fractional hypoxic volume, HV hypoxic volume

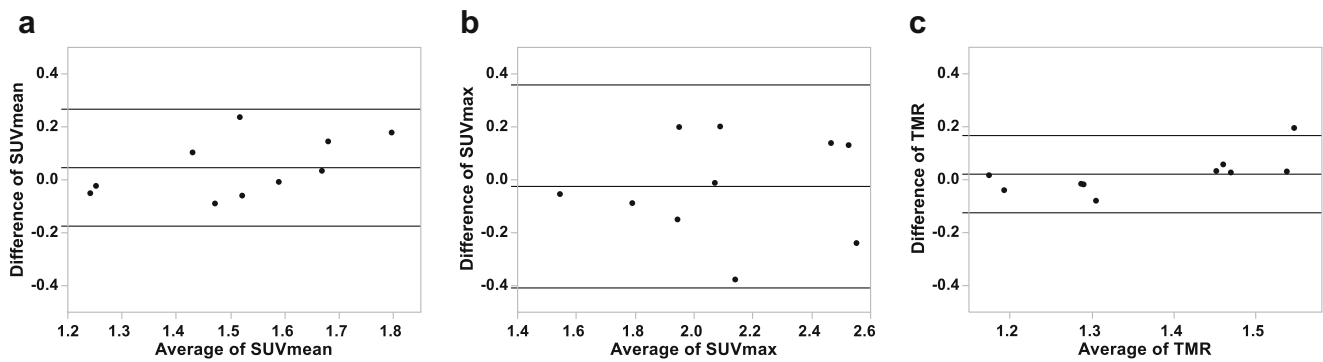


Fig. 2 Bland-Altman plots of SUVmean (a), SUVmax (b), and tumour-to-muscle uptake ratio (TMR) (c) of repeated [^{18}F]EF5 PET/CT scans. Beginning from the most superior one, the three solid lines represent the

upper limit of agreement (LoA), the mean difference, and the lower LoA, respectively

In the voxel-by-voxel analysis, the mean of Pearson correlation coefficients between TMR of [^{18}F]EF5 in the repeated scans within individual patients was 0.65 (range 0.48–0.87). The scatterplots of individual patients are presented in Fig. 3. The mean differences of voxel-level TMRs of individual patients between the paired scans with upper and lower LoA are shown in Table 3. The mean calculated from mean differences of individual patients was 0.02 ± 0.07 . For the pooled dataset, the mean difference of voxelwise TMR was 0.03 ± 0.20 , with an upper and lower LoA of 0.41 and -0.36 , respectively (Fig. 4), and the absolute CoR and relative CoR were 0.39 and 28%, respectively.

Discussion

This study was designed to assess the repeatability of [^{18}F]EF5 PET/CT among HNC patients before the onset of definitive CRT. For each individual oncologic PET tracer, a test-retest study is a fundamental part of the feasibility evaluation for clinical applications [16]. However, from a methodological point of view, the comparison of repeated hypoxia PET scans has been considered difficult to be reduced into a robust assessment of accuracy in quantitative analysis since the measured phenomenon has been assumed to change over the short term due to the so-called cycling hypoxia [17]. Thus, the results of hypoxia PET repeatability studies have been interpreted as a combination of the technical repeatability of measurements and true changes in tumour oxygenation. Nevertheless, at the resolution reported here, which is entirely adequate to apply image-guided targeted therapy, cycling hypoxia does not appear to be a significant factor.

In this study, the [^{18}F]EF5 uptake in primary HNC was comparable to those reported previously [7, 8]. The definition of tumour VOI for repeatability assessment was based on MATV in the [^{18}F]FDG image to facilitate comparison with some of the previous clinical hypoxia PET repeatability studies [9, 11]. For this purpose, we prefer MATV to pure CT-based delineation in

order to avoid the inclusion of background tissues, which might cause an erroneously high repeatability of intratumour [^{18}F]EF5 uptake [18, 19]. This is mainly due to the presumption that tissues in the periphery with low [^{18}F]FDG uptake would be expected to show less dynamic and more repeatable hypoxia tracer uptake than the core of the tumour. Posterior neck muscles were used as reference tissue and the uptake of [^{18}F]EF5 in muscle within paired scans was observed to have a good correlation (ICC 0.84) and high repeatability with a relative CoR of 10%. Furthermore, the findings in venous blood samples supported the assumption of stable radioactivity concentration in the blood pool, and thus, highly repeatable activity measurements.

A TMR of 1.5 was derived as a threshold representing hypoxic tissue from a previous study in HNC, where this level was determined based on a voxelwise comparison between perfusion and uptake of [^{18}F]EF5 in tumour [7]. The median FHV of the tumours (20.2%) in the present study was slightly smaller compared to some previous studies with [^{18}F]FMISO and [^{18}F]FETNIM PET/CT using MATV-based GTV delineation in HNC patients [19]. However, any threshold for hypoxia is an estimation reflecting the present patient population and applied methodology and instrumentation. Therefore, correlation and agreement between the repeated scans across the whole scale of uptake rates including “non-hypoxic” areas are essential for correct judgement in a test-retest study.

We found that tumour-level parameters (SUVmean, SUVmax, and TMR) showed a high correlation and repeatability between the paired [^{18}F]EF5 scans. These results agree with the previous studies in HNC and lung cancer using hypoxia-activated 2-nitroimidazole tracers [^{18}F]FMISO [10, 11] and [^{18}F]HX4 [12]. However, the oldest study using [^{18}F]FMISO PET/CT [9] reported a lower correlation between the repeated scans compared to ours and the above-mentioned three studies [10–12]. Several explanations for this controversy have been proposed, including an inconsistent uptake time within the repeated scans, the use of either 2D or 3D acquisition modes, and the variability of image co-registration algorithms [11]. The results of the present study support the

Fig. 3 Scatterplots of voxelwise tumour-to-muscle uptake ratios (TMR). The X-axis represents the first and the Y-axis the second of the [^{18}F]EF5 PET/CT scans. Solid lines indicate the cutoff level for hypoxia (TMR 1.5)

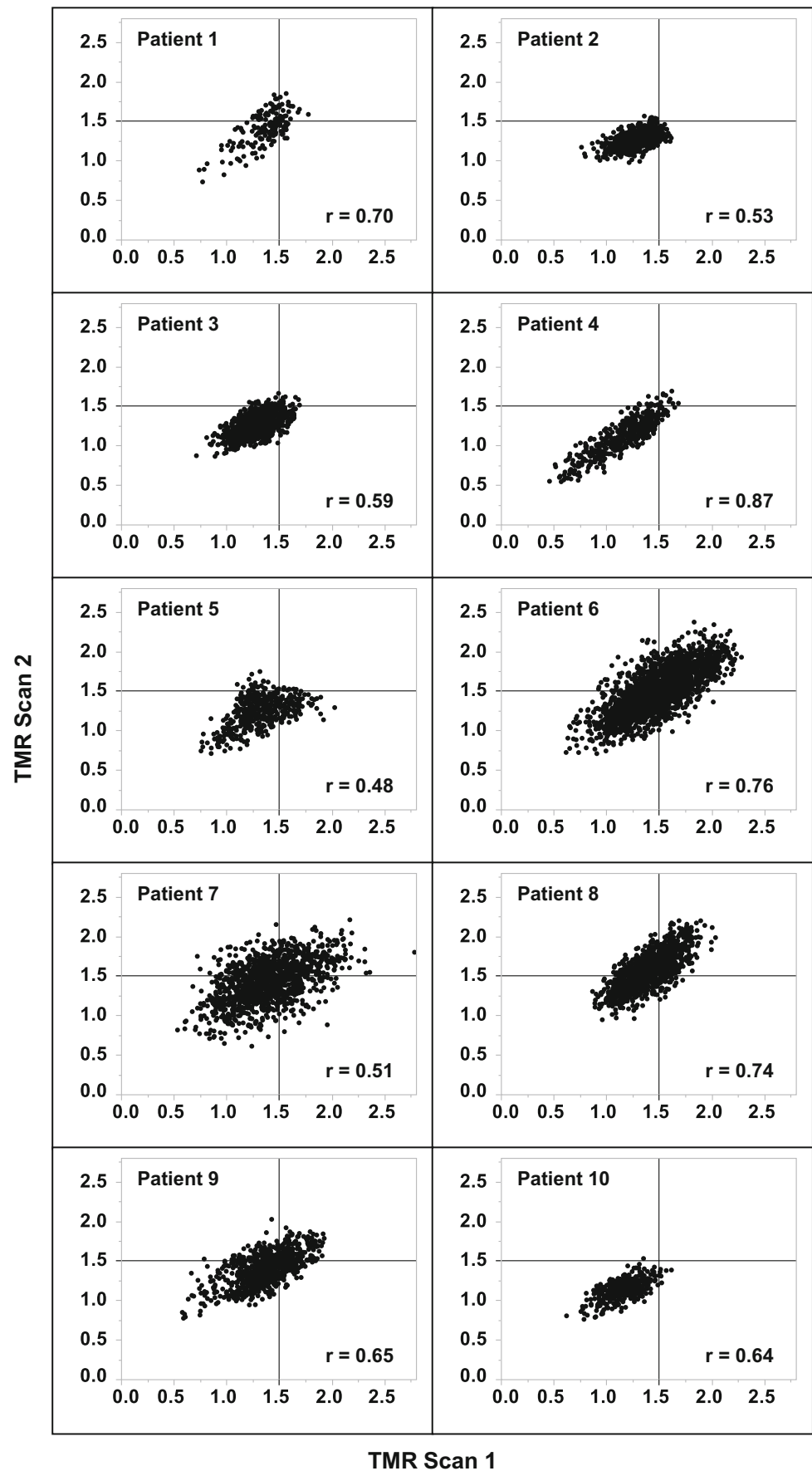


Table 3 Results of voxel-level agreement analysis between tumour-to-muscle uptake ratios of repeated [¹⁸F]EF5 PET/CT scans

Patient nr	Number of voxels	Mean ± SD difference (95% CI)	Upper LoA	Lower LoA
1	165	0.05 ± 0.16 (0.02 — 0.07)	0.36	−0.26
2	757	0.00 ± 0.12 (−0.01 — 0.01)	0.23	−0.23
3	1159	−0.02 ± 0.13 (−0.02 — −0.01)	0.23	−0.26
4	470	−0.03 ± 0.12 (−0.04 — −0.02)	0.21	−0.27
5	476	−0.07 ± 0.21 (−0.09 — −0.05)	0.34	−0.48
6	2306	0.02 ± 0.20 (0.01 — 0.02)	0.40	−0.37
7	1233	0.03 ± 0.28 (0.02 — 0.05)	0.59	−0.52
8	1155	0.18 ± 0.16 (0.17 — 0.19)	0.48	−0.13
9	860	0.03 ± 0.18 (0.02 — 0.04)	0.39	−0.33
10	394	−0.02 ± 0.12 (−0.03 — 0.00)	0.22	−0.26
Pooled dataset	8975	0.03 ± 0.20 (0.02 — 0.03)	0.41	−0.36

CI confidence interval, LoA limit of agreement

perception of the need for highly consistent imaging protocol and data processing algorithms to be applied in analyses addressing the repeatability of hypoxia PET studies.

The voxel-by-voxel analysis showed good or moderate spatial correlation in [¹⁸F]EF5 uptake between the paired scans. Using $r > 0.5$ as a level of strong correlation and reproducible results similarly as in two previous studies [9, 12], we observed a strong voxelwise correlation for 9 out of 10 of our patients. On the other hand, a slightly lower agreement using relative mean difference and relative CoR between the voxel uptake of paired scans was observed in this study compared to those of Grkovski et al. [11] and Zegers et al. [12]. Nevertheless, comparing results between the present and all previous [9–12] studies is challenging due to some heterogeneities in acquisition parameters. A crucial parameter affecting voxel-level repeatability is the used voxel size in images which was not reported by Okamoto et al. [10] and Zegers et al. [12]. We used a voxel size of 3.65 × 3.65 × 3.27 mm,

similar to our institutional diagnostic protocol and representing the high end of the reported resolutions used in other corresponding studies [9, 11]. Another parameter not uniformly available for comparison is the tumour size [10], which has an influence on partial volume effect. However, the distribution of tumour size in our study seems to be comparable to those reported in previous studies of HNC [9, 12]. Finally, special attention should be given to statistical methods to address repeatability in a test-retest design, and following this, we calculated both correlation and agreement values for all uptake parameters [20].

There is clearly a trend of larger variability in tumour hypoxia when a longer time period is assessed, although limited data on temporal changes is available with hypoxia detection methods other than PET [21]. Our previous preclinical PET/CT study showed a large variation in intratumour uptake of [¹⁸F]EF5 in xenografted HNC at different stages of tumour growth up to 5–36 days apart [22]. On the other hand, preclinical studies using a short interval from 6 h to 1 day have reported a high repeatability of hypoxia imaging with [¹⁸F]FAZA PET [23] and [¹⁸F]FMISO PET [24]. In the present study, the median time of 7 days between the repeated scans was longer than that in several previous clinical studies that reported an average interval of 1–3 days [9–12, 25]. However, we did not observe any trend for lower repeatability of tracer uptake parameters compared to those studies where interscan time was shorter. Consequently, we consider that the impact of the difference between the intervals of the present and previous studies is small.

This study also has some limitations. In line with previous studies [9–12], the number of patients was reasonably small. In addition, all of the study subjects were men. These limitations derived from the challenges in subject enrollment as well as the remarkably higher incidence of pharyngeal cancer among male patients. Partial volume effect might increase the variability of tracer uptake between paired scans, especially among small and irregularly shaped tumours. Finally, the repeated setup of patients

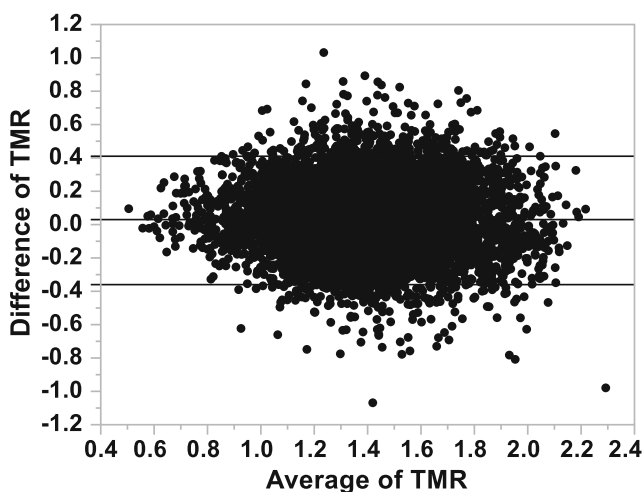


Fig. 4 A Bland-Altman plot of voxelwise tumour-to-muscle uptake ratios (TMRs) from the pooled data of all patients. Beginning from the most superior one, the three solid lines represent the upper limit of agreement (LoA), the mean difference, and the lower LoA, respectively

as well as the co-registration of images are known to be prone to some kind of inaccuracies, although optimal methods and head and neck immobilisation masks are used [26]. Nevertheless, these issues that may cause a decrease in the repeatability of tracer uptake are pragmatic challenges, present in everyday clinical imaging and image analysis.

The feasibility of [^{18}F]EF5 PET/CT for guiding RT dose escalation or adaptation deserves attention in the future. The pretreatment hypoxia-specific signal of [^{18}F]EF5 at 3 h from injection in HNC is repeatable and comparable to those of [^{18}F]FMISO and [^{18}F]HX4 at 4 and 2 h, respectively [10, 12]. Given the similarities in the chemical properties of the three tracers, this is not surprising, while small differences in their sensitivities to acute vs. chronic hypoxia may occur [27]. Recently, a few clinical studies were performed where changes in tumour hypoxia defined with PET/CT were monitored during the first weeks of RT. In general, these studies state that residual hypoxia after the first week or two of RT is more stable and shows more prognostic significance compared to pretreatment tumour hypoxia [28, 29]. Thus, an important future study should investigate the stability and prognostic significance of [^{18}F]EF5 uptake during the course of RT. Another central clinical point of view to be assessed is whether [^{18}F]EF5 PET/CT qualifies for the selection of patients for hypoxia-targeted interventions, such as treatment with hypoxia-avid radiosensitisers or hypoxia-activated prodrugs [30].

Conclusion

A high repeatability of tumour-level tracer uptake was observed in the paired [^{18}F]EF5 PET/CT scans acquired before the onset of CRT. The voxel-by-voxel analysis showed predominantly good correlation and agreement between the repeated scans. We thus encourage further evaluation of [^{18}F]EF5 PET/CT for guiding hypoxia-targeted treatment interventions.

Acknowledgements We thank the staff of the Turku PET Centre, the Department of Oncology and Radiotherapy, and the Department of Otorhinolaryngology – Head and Neck Surgery for their excellent help in several steps of this project. Financial support was provided by the Hospital District of Southwest Finland, the National Graduate School of Clinical Investigation, the Cancer Society of Finland, the Finnish Medical Foundation, and the Foundation of the Finnish Otorhinolaryngological Society.

Compliance with ethical standards

Conflict of interest All authors declare that they have no conflicts of interest.

Ethical approval 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 article does not contain any studies with animals performed by any of the authors.

Informed consent Informed consent was obtained from all individual participants included in the study.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Barker HE, Paget JT, Khan AA, Harrington KJ. The tumour micro-environment after radiotherapy: mechanisms of resistance and recurrence. *Nat Rev Cancer*. 2015;15:409–25.
- Vaupel P, Mayer A. Hypoxia in cancer: significance and impact on clinical outcome. *Cancer Metastasis Rev*. 2007;26:225–39.
- Bredell MG, Ernst J, El-Kochairi I, Dahlem Y, Ikenberg K, Schumann DM. Current relevance of hypoxia in head and neck cancer. *Oncotarget*. 2016;7:50781–804.
- Welz S, Mönnich D, Pfannenbergs C, Nikolau K, Reimold M, La Fougère C, et al. Prognostic value of dynamic hypoxia PET in head and neck cancer: results from a planned interim analysis of a randomized phase II hypoxia-image guided dose escalation trial. *Radiother Oncol*. 2017;124:526–32.
- Koch CJ. Measurement of absolute oxygen levels in cells and tissues using oxygen sensors and 2-nitroimidazole EF5. *Methods Enzymol*. 2002;352:3–31.
- Koch CJ, Evans SM. Optimizing hypoxia detection and treatment strategies. *Semin Nucl Med*. 2015;45:163–76.
- Komar G, Seppänen M, Eskola O, Lindholm P, Grönroos TJ, Forsback S, et al. ^{18}F -EF5: a new PET tracer for imaging hypoxia in head and neck cancer. *J Nucl Med*. 2008;49:1944–51.
- Komar G, Lehtiö K, Seppänen M, Eskola O, Levola H, Lindholm P, et al. Prognostic value of tumour blood flow, [^{18}F]EF5 and [^{18}F]FDG PET/CT imaging in patients with head and neck cancer treated with radiochemotherapy. *Eur J Nucl Med Mol Imaging*. 2014;2042–50.
- Nehmeh SA, Lee NY, Schröder H, Squire O, Zanzonico PB, Erdi YE, et al. Reproducibility of intratumor distribution of ^{18}F -fluoromisonidazole in head and neck cancer. *Int J Radiat Oncol Biol Phys*. 2008;70:235–42.
- Okamoto S, Shiga T, Yasuda K, Ito YM, Magota K, Kasai K, et al. High reproducibility of tumor hypoxia evaluated by ^{18}F -Fluoromisonidazole PET for head and neck cancer. *J Nucl Med*. 2013;54:201–7.
- Grkovski M, Schwartz J, Rimner A, Schöder H, Carlin SD, Zanzonico PB, et al. Reproducibility of ^{18}F -fluoromisonidazole intratumour distribution in non-small cell lung cancer. *EJNMMI Res*. 2016;6:79.
- Zegers CML, van Elmp W, Szardenings K, Kolb H, Waxman A, Subramaniam RM, et al. Repeatability of hypoxia PET imaging using [^{18}F]HX4 in lung and head and neck cancer patients: a prospective multicenter trial. *Eur J Nucl Med Mol Imaging*. 2015;42:1840–9.
- Blitzer GC, Smith MA, Harris SL, Kimple RJ. Review of the clinical and biologic aspects of human papillomavirus-positive squamous cell carcinomas of the head and neck. *Int J Radiat Oncol Biol Phys*. 2014;88:761–70.

14. Eskola O, Grönroos TJ, Forsback S, Tuomela J, Komar G, Bergman J, et al. Tracer level electrophilic synthesis and pharmacokinetics of the hypoxia tracer [^{18}F]EF5. *Mol Imaging Biol.* 2012;14:205–12.
15. Minn H, Suilamo S, Seppälä J. Impact of PET/CT on planning of radiotherapy in head and neck cancer. *Q J Nucl Med Mol Imaging.* 2010;54:521–32.
16. Lodge MA. Repeatability of SUV in oncologic ^{18}F -FDG PET. *J Nucl Med.* 2017;58:523–32.
17. Dewhirst MW, Cao Y, Moeller B. Cycling hypoxia and free radicals regulate angiogenesis and radiotherapy response. *Nat Rev Cancer.* 2008;8:425–37.
18. Hoeben BA, Bussink J, Troost EG, Oyen WJ, Kaanders JH. Molecular PET imaging for biology-guided adaptive radiotherapy of head and neck cancer. *Acta Oncol.* 2013;52:1257–71.
19. Chirla R, Marcu LG. PET-based quantification of statistical properties of hypoxic tumor subvolumes in head and neck cancer. *Phys Med.* 2016;32:23–35.
20. Vaz S, Falkmer T, Passmore AE, Parsons R, Andreou P. The case for using the repeatability coefficient when calculating test-retest reliability. *PLoS One.* 2013;8(9):e73990.
21. Kelada OJ, Carlson DJ. Molecular imaging of tumor hypoxia with positron emission tomography. *Radiat Res.* 2014;181:335–49.
22. Silvonien A, Silén J, Forsback S, Löyttyniemi E, Schrey AR, Solin O, et al. Evaluation of repeated [^{18}F]EF5 PET/CT scans and tumor growth rate in experimental head and neck carcinomas. *EJNMMI Res.* 2014;4:65.
23. Busk M, Mortensen LS, Nordmark M, Overgaard J, Jakobsen S, Hansen KV, et al. PET hypoxia imaging with FAZA: reproducibility at baseline and during fractionated radiotherapy in tumor-bearing mice. *Eur J Nucl Med Mol Imaging.* 2013;40:186–97.
24. Whisenant JG, Peterson TE, Fluckiger JU, Tantawy MN, Ayers GD, Yankeelov TE. Reproducibility of static and dynamic ^{18}F -FDG, ^{18}F -FLT and ^{18}F -FMISO MicroPET studies in a murine model of HER2+ breast cancer. *Mol Imaging Biol.* 2013;15:87–96.
25. Yue J, Yang Y, Cabrera AR, Sun X, Zhao S, Xie P, et al. Measuring tumor hypoxia with ^{18}F -FETNIM PET in esophageal squamous cell carcinoma: a pilot clinical study. *Dis Esophagus.* 2012;25:54–61.
26. Park ET, Park SK. Setup uncertainties for inter-fractional head and neck cancer in radiotherapy. *Oncotarget.* 2016;7:46662–7.
27. Peeters SG, Zegers CM, Lieuwes NG, van Elmpt W, Eriksson J, van Dongen GA, et al. A comparative study of the hypoxia PET tracers [^{18}F]HX4, [^{18}F]FAZA, and [^{18}F]FMISO in a preclinical tumor model. *Int J Radiat Oncol Biol Phys.* 2015;91:351–9.
28. Bollineni VR, Koole MJ, Pruijm J, Brouwer CL, Wiegman EM, Groen HJ, et al. Dynamics of tumor hypoxia assessed by ^{18}F -FAZA PET/CT in head and neck and lung cancer patients during chemoradiation: possible implications for radiotherapy treatment planning strategies. *Radiother Oncol.* 2014;113:198–203.
29. Zips D, Zöphel K, Abolmaali N, Perrin R, Abramyuk A, Haase R, et al. Exploratory prospective trial of hypoxia-specific PET imaging during radiochemotherapy in patients with locally advanced head-and-neck cancer. *Radiother Oncol.* 2012;105:21–8.
30. Rischin D, Hicks RJ, Fisher R, Binns D, Corry J, Porceddu S, et al. Trans-Tasman radiation oncology group study 98.02. Prognostic significance of [^{18}F]misonidazole positron emission tomography-detected tumor hypoxia in patients with advanced head and neck cancer randomly assigned to chemoradiation with or without tirapazamine: a substudy of trans-Tasman radiation oncology group study 98.02. *J Clin Oncol.* 2006;24:2098–104.