Repeatability of Quantitative ^{18}F-Fluoromethylcholine PET/CT Studies in Prostate Cancer

Daniela E. Oprea-Lager[1], Gem Kramer[1], Peter M. van de Ven[2], Alfons J.M. van den Eertwegh[3], Reindert J.A. van Moorselaar[4], Patrick Schober[5], Otto S. Hoekstra[1], Adriaan A. Lammertsma[1], and Ronald Boellaard[1]

[1]Department of Radiology and Nuclear Medicine, VU University Medical Center, Amsterdam, The Netherlands; [2]Department of Epidemiology and Biostatistics, VU University Medical Center, Amsterdam, The Netherlands; [3]Department of Medical Oncology, VU University Medical Center, Amsterdam, The Netherlands; [4]Department of Urology, VU University Medical Center, Amsterdam, The Netherlands; and [5]Department of Anesthesiology, VU University Medical Center, Amsterdam, The Netherlands

Repeatable quantification is essential when using ^{18}F-fluoromethylcholine PET/CT to monitor treatment response in prostate cancer. It has been shown that SUV normalized to the area under the blood activity concentration curve (SUV_{AUC}) provides a better correlation with full kinetic analysis than does standard SUV. However, the precision of SUV_{AUC} is not known yet. The purpose of this study was to assess the repeatability of various semiquantitative ^{18}F-fluoromethylcholine parameters in prostate cancer. **Methods:** Twelve patients (mean age ± SD, 64 ± 8 y) with metastasized prostate cancer underwent two sets of ^{18}F-fluoromethylcholine PET/CT scans, on consecutive days. Each set consisted of a 30-min dynamic PET/CT scan of the chest after intravenous administration of 200 MBq of ^{18}F-fluoromethylcholine, followed by a whole-body PET/CT scan at 40 min. The dynamic scan was used to derive the area under the blood activity concentration curve. Lesion uptake was derived from the whole-body scan using various types of volumes of interest: maximum, peak, and mean. Each of these parameters was normalized to injected activity per body weight, area under the blood activity concentration curve, and blood concentration itself at 40 min, resulting in several types of SUVs: SUV, SUV_{AUC}, and SUV_{TBR}. The test–retest repeatability of these metrics, as well as metabolic tumor volume (MTV) and total uptake of choline in the lesion, were studied. The level of agreement between test–retest data and reliability was assessed using Bland–Altman plots, repeatability coefficients, and intraclass correlation coefficients (ICCs). **Results:** A total of 67 choline-avid metastases were identified: 44 bone lesions and 23 lymph node lesions. In the case of SUV_{max}, the repeatability coefficients for SUV, SUV_{AUC}, and SUV_{TBR} were 26% (ICC, 0.95), 31% (ICC, 0.95), and 46% (ICC, 0.89), respectively. Similar values were obtained for SUV_{peak} and SUV_{mean}. The repeatability of SUV_{AUC} was comparable to that of SUV_{max}, SUV_{peak}, and SUV_{mean}. Tissue type and tumor localization did not affect repeatability. An MTV of less than 4.2 cm^3 had larger variability than larger volumes (repeatability coefficient, 45% vs. 29%; $P = 0.048$). The repeatability coefficient did not significantly differ between lesions with SUV_{peak} above or below the median value of 8.3 (19% vs. 28%; $P = 0.264$). **Conclusion:** The repeatability of SUV_{AUC} was comparable to that of standard SUV. The repeatability coefficients of various semiquantitative ^{18}F-fluoromethylcholine parameters (SUV, MTV, and total uptake in the lesion) were approximately 35%. Larger differences are likely to represent treatment effects.

Received Oct. 2, 2015; revision accepted Dec. 1, 2015.

For correspondence or reprints contact: Daniela E. Oprea-Lager, Department of Radiology and Nuclear Medicine, VU University Medical Center, P.O. Box 7057, 1007 MB Amsterdam, The Netherlands.

E-mail: d.oprea-lager@vumc.nl

Published online Dec. 23, 2015.

Key Words: repeatability; positron emission tomography (PET); prostate cancer; choline; standardized uptake value (SUV)

J Nucl Med 2016; 57:721–727
DOI: 10.2967/jnumed.115.167692

Prostate cancer is the second most common cancer in men worldwide and was the third most diagnosed malignancy in Europe in 2012, with 92,000 deaths (*1,2*). This androgen-dependent neoplasm is characterized by a good initial response to antihormonal therapy and an unpredictable latent castration-resistant status (*3*). At the beginning of this decennium, molecular profiling studies improved our knowledge about the heterogeneous biologic behavior of prostate cancer. It was found that even in the presence of a castrate range of serum testosterone (<1.7 nmol/L), a proportion of tumors in castration-resistant prostate cancer patients remains dependent on androgen-receptor signaling for growth (*4*). Five potential mechanisms of development of castration-resistant prostate cancer were described, based on ligand and androgen-receptor dependence (*5*).

Today, several therapeutic options against castration-resistant prostate cancer prevail, including cytotoxic (docetaxel, cabazitaxel), antihormonal (abiraterone, enzalutamide), immunotherapeutic (sipuleucel-T), and bone-targeting (^{223}Ra-dichloride) agents (*6–13*). However, despite this variety of new agents with demonstrated improvement in life expectancy, the proper sequencing (e.g., modality and timing) in individual patients with metastatic prostate cancer is unclear (*14*). Because therapeutic options vary greatly with the stage and grade of the disease, the specific pattern of metastatic spread (i.e., hematogenous or lymphatic), and the dominant phenotype, accurate diagnostic instruments for response evaluation are essential (*15,16*).

Noninvasive hybrid PET/CT is a valuable diagnostic tool by virtue of its combining metabolic and anatomic information in vivo (*17*). Encouraging results have been reported on the usefulness of radiolabeled-choline PET/CT in prostate cancer (*18,19*). Apart from its main recognized application in restaging disease in cases of biochemical relapse (*20,21*), ^{18}F-fluoromethylcholine might also qualify as a biomarker of response to therapy. Because, conceptually, choline uptake represents viable tumor cells, tracer uptake changes over time might serve as an improved readout of treatment efficacy.

In vitro experiments have shown promising results for the use of radiolabeled choline to monitor antiandrogen treatment and chemotherapy (*22,23*). Recently, simplified quantitative methods for measuring

^{18}F-fluoromethylcholine uptake have been developed and validated (*24*); SUV normalized to the area under the blood activity concentration curve (SUV_{AUC}) correlates better with full kinetic analysis than does standard SUV (*24*). However, the precision of SUV_{AUC}, and also that of SUV itself, is as yet unknown.

The purpose of this study was to prospectively assess the repeatability of semiquantitative ^{18}F-fluoromethylcholine PET/CT parameters in prostate cancer, also including metabolic tumor volume (MTV) and total uptake of choline in the lesion. Such knowledge is essential for proper interpretation of changes in the ^{18}F-fluoromethylcholine signal over time, thus improving personalized therapy strategies for prostate cancer patients.

MATERIALS AND METHODS

Patients

Twelve patients with histologically proven prostate cancer (4 of whom had castration-resistant prostate cancer) and lymphatic or hematogenous metastases were included prospectively. The inclusion criteria required at least 2 metastases (diameter ≥ 1.5 cm) detected by conventional imaging performed no more than 3 mo before PET/CT and the ability to remain supine for 60 min. The exclusion criteria were claustrophobia and coexistence of multiple malignancies. The study was approved by the Medical Ethics Review Committee of VU University Medical Center. Before inclusion, each patient signed a written informed consent form after receiving a verbal and written explanation of the study.

Personal and demographic data were collected (age, body weight, Gleason score, prostate-specific antigen level (ng/mL) at the time of PET/CT, and information on previous therapy), as well as the characteristics of the metastatic lesions (location [intrathoracic, intraabdominal, or pelvic], number, and type [bone or lymph node]). Values are presented as mean ± SD.

Data Acquisition

All patients underwent ^{18}F-fluoromethylcholine PET/CT on 2 consecutive days. The tracer was synthesized according to a previously published method (*24*). The minimal interval between the last treatment and the first PET/CT scan was 19 d. Patient preparation was similar to that required for ^{18}F-FDG (*25*). The patients were scanned using a Gemini TF-64 PET/CT scanner (Philips).

Each patient underwent a low-dose CT scan (30 mAs, 120 kV) followed by a 30-min dynamic PET scan of the chest (*24*), centered over a large blood-pool structure (e.g., the ascending aorta), to obtain an image-derived input function. At the start of the dynamic ^{18}F-fluoromethylcholine scan, a bolus injection of 205 ± 9 MBq (day 1) and 206 ± 7 MBq (day 2) of ^{18}F-fluoromethylcholine was administered intravenously using an automated injector (Medrad), which was flushed with 40 mL of saline (5 mL at 0.8 mL·s^{-1} followed by 35 mL at 2 mL·s^{-1}). Dynamic PET data were normalized and corrected for decay, scatter, random coincidences, and photon attenuation and were reconstructed into 25 frames (1 × 10, 8 × 5, 5 × 20, 5 × 60, 3 × 150, and 3 × 300 s) with a matrix size of 144 × 144 × 45 and voxels of 4 × 4 × 4 mm^3, using a 3-dimensional row action maximum likelihood algorithm.

After a standard (5-min) micturition break to warrant a proper visual assessment of the pelvic region, a whole-body ^{18}F-fluoromethylcholine scan (mid-thigh to skull base) was performed 40 min after injection. After this PET acquisition (10 bed positions at 2 min each), a second low-dose CT scan (50 mAs, 120 kV) was acquired for anatomic correlation and attenuation correction. Whole-body data were corrected for dead time, decay, scatter, and randoms and were reconstructed into 34 frames (1 × 10, 8 × 5, 4 × 10, 3 × 20, 5 × 30, 5 × 60, 4 × 150, and 4 × 300 s) with a matrix size of 144 × 144 voxels (4 × 4 × 4 mm^3), using iterative reconstruction (binary large-object ordered-subsets time-of-fight). The transaxial spatial resolution was about 5 mm full width at half maximum in the center of the field of view, similar to that of the dynamic scan.

Data Analysis

The reconstructed images were transferred to off-line workstations for further analysis. Data were analyzed on a volume-of-interest (VOI) basis (*26*). The dynamic scan was used to derive the area under the blood activity concentration curve, by defining a cylindric VOI 1.5 cm in diameter extending over 5 consecutive axial planes within the lumen of the ascending aorta. Next, lesion uptake (defined as tracer accumulation exceeding local background activity and incompatible with physiologic tracer biodistribution) was derived from the whole-body scan.

The term *metabolic tumor volume* was used to indicate volumes that were derived directly from the PET studies, quantified as the VOI size (*26*). VOIs were defined by a semiautomatic delineation tool, applying a background-adapted 50% of maximum isocontour (i.e., contour value = 50% of maximum + background). SUV_{max}, SUV_{peak}, and SUV_{mean} were calculated for each VOI. In addition, each of these parameters was normalized to injected activity per body weight (SUV), area under the blood activity concentration curve (SUV_{AUC}) as derived from the dynamic scan, and blood concentration at 40 min after injection (SUV tumor-to-background ratio [SUV_{TBR}]) as derived from the whole-body scan. For the parameters associated with SUV_{mean}, total choline uptake in the lesion (defined as $SUV_{mean} \times MTV$) was also calculated, resulting in total uptake, area under the total uptake curve, and total uptake tumor-to-background ratio. The test–retest repeatability of all metrics was calculated using both standard repeatability coefficients (1.96 × SD of difference between test and retest) and relative repeatability coefficients (percentage test–retest difference).

Statistical Analysis

Data were analyzed using SPSS, version 15.0 (IBM). Test–retest repeatability was quantified using intraclass correlation coefficients (ICCs; based on absolute agreement) and repeatability coefficients and was displayed graphically using Bland–Altman or box plots. ICCs were calculated for each SUV measure as the ratio of between-lesion variability and total variability (both between-lesion and within-lesion). Variance components for patient, lesion within patient, and repeated measurements within lesion were estimated using mixed models. Total variability was calculated as the sum of these 3 variance components. Between-lesion variability was calculated as the sum of the variance components for patient and for lesion within patient. Confidence intervals for ICCs were determined using the delta method. Reliability coefficients were based on the relative difference:

$$100\% \times (SUV_{test} - SUV_{retest})/[0.5 \times (SUV_{test} + SUV_{retest})].$$

When repeatability coefficients were calculated, the method of Bland and Altman (*27*) was used to take into account correlation of measurements between different lesions within the same patient. Differences in the repeatability of the different SUV measures (SUV_{max}, SUV_{peak}, and SUV_{mean}) and methods of normalization (SUV, SUV_{AUC}, and SUV_{TBR}) were assessed by comparing the variances of the relative test–retest differences, using the Pitman–Morgan test (*28*) for correlated variances. Differences in the repeatability of SUV measures related to location, lesion type, and lesion size were assessed by comparing the variances of the relative test–retest differences between lesion subgroups, using the Levene test (*29*). A Bonferroni correction was used to control overall type I error to 5% within each set of comparisons.

RESULTS

Patients

The 12 patients had a mean age of 64 ± 8 y, a mean body weight of 88 ± 9 kg, a Gleason score of 7 ($n = 3$) or higher ($n = 9$), and a median prostate-specific antigen level of 46 ng/mL (range, 2–226 ng/mL) at the time of PET/CT (Table 1). All patients had been treated previously by antihormone therapy, with the addition of chemotherapy alone in 3, chemotherapy plus external-beam radiotherapy in 1, external-beam radiotherapy alone in 2, external-beam radiotherapy plus prostatectomy in 2, external-beam radiotherapy plus immunotherapy in 1, and prostatectomy plus lymph node dissection in 1.

At PET, 67 metastases were identified (median, 6 per patient; range, 3–8), 44 of which were to bone and 23 to lymph node metastases. Twelve metastases were above the diaphragm; the other 55 were intraabdominal or pelvic. The malignant nature of the metastases was confirmed radiologically on the basis of progression of preexistent lesions or new metastatic sites. Median size and interquartile range were 4.9 cm^3 and 7.6 cm^3, respectively, for VOI and 8.3 and 5.2, respectively, for SUV_{peak}. Repeatability data were analyzed using a volume threshold of 4.2 cm^3 based on a repeatability study of MTV with ^{18}F-FDG and ^{18}F-fluorothymidine in lung cancer (*26*). In that study, changes in ^{18}F-FDG uptake of more than 37% for lesions larger than 4.2 cm^3 were found to represent a biologic effect. This volume threshold corresponds by approximation to a diameter of 2 cm (for spheric metastatic lesions), which equals about 4 times the spatial resolution of PET, below which it is not possible to determine quantification, VOI definition, or detectability without interference by partial-volume effects.

Repeatability of Semiquantitative Parameters

The repeatability of each semiquantitative parameter (i.e., SUV, MTV, and total lesion uptake) was studied as a function of uptake (median SUV_{peak} of 8.3), MTV (larger or smaller than 4.2 cm^3), metastatic tissue type (bone or lymph node), and location (intrathoracic vs. abdominal or pelvic).

Repeatability of SUV. In the case of SUV_{max}, the repeatability coefficients (for relative differences) were 26% for SUV, 31% for SUV_{AUC}, and 46% for SUV_{TBR}. Similar values were observed for SUV_{peak} and SUV_{mean}. All ICCs were 0.95 or 0.96 for SUV and SUV_{AUC} and 0.89 for SUV_{TBR} (Table 2).

Nine pairwise comparisons for different methods of SUV normalization were performed to estimate variance in relative test–retest differences. After correction for multiple comparisons (Bonferroni-corrected significant difference, $P < 0.0056$ at the 5% level and $P < 0.0011$ at the 1% level), only the SUV_{TBR} parameters were found to have consistently larger variances ($P < 0.001$) (Supplemental Table 1). The relative percentage differences between test–retest data and the means for SUV_{max}, SUV_{peak}, and SUV_{mean} and their normalizations are presented in Figure 1. The repeatability of SUV_{AUC} for SUV_{max}, SUV_{peak}, and SUV_{mean} was comparable to that of the corresponding SUV measures (Supplemental Table 2).

Repeatability coefficients did not differ between lesions with an SUV_{peak} above or below the median value of 8.3 (19% vs. 28%; $P = 0.264$) (Fig. 2 and Supplemental Fig. 1). These values were comparable to those of $SUV_{peak/AUC}$ (23% [$SUV_{peak} > 8.3$] vs. 31% [$SUV_{peak} < 8.3$]; $P = 0.136$). Moreover, the repeatability of SUV_{peak} and SUV_{mean} was independent of MTV (Fig. 3), as well as of tissue type and tumor location (Supplemental Fig. 2; Supplemental Table 3).

Repeatability of MTV. For MTV, the test–retest difference was 0.03 ± 1.63 and the relative test–retest difference 36% (Fig. 4). The repeatability coefficients for MTV were independent of SUV_{peak} (34.25% [$SUV_{peak} > 8.3$] vs. 36.43% [$SUV_{peak} < 8.3$]; $P = 0.933$]. An MTV of less than 4.2 cm^3 had larger variability than larger volumes (repeatability coefficient, 45% vs. 29%; $P = 0.048$) (Supplemental Fig. 1). The repeatability coefficients for MTV did not differ between bone/lymph nodes (34% vs. 36.4%;

TABLE 1
Patient Characteristics

Patient no.	Age (y)	Weight (kg)	Gleason score		PSA (ng/mL)	Previous therapy					
			7	>7		AT	CT	EBRT	RALP	LND	IT
1	65	78		9	93	1	1				
2	55	89		8	18	1	1	1			
3	74	90	7		43	1			1	1	
4	68	79		10	8	1	1				
5	66	83		8	54	1		1	1		
6	60	85		9	137	1		1	1		
7	54	98	7		2	1		1			
8	72	84	7		54	1		1			
9	54	85		9	2	1					
10	54	108		9	226	1		1			1
11	71	81		8	39	1	1				
12	70	101		8	49	1					

PSA = prostate-specific antigen level at time of PET/CT; AT = antihormone therapy; CT = chemotherapy; EBRT = external-beam radiotherapy; RALP = robot-assisted laparoscopic prostatectomy; LND = lymph node dissection; IT = immunotherapy.

TABLE 2
Test–Retest Differences, ICCs, and Repeatability Coefficients for Different Types of SUVs and Their Normalizations

SUV type	Normalization	TRT diff*	ICC	RC†
SUV_{max}	SUV	−0.38 ± 1.45	0.95 (0.91–0.98)	26.1%
	SUV_{AUC}	0 ± 0.88	0.95 (0.92–0.99)	30.9%
	SUV_{TBR}	1.0 ± 3.9	0.89 (0.81–0.97)	46.2%
SUV_{peak}	SUV	−0.26 ± 0.97	0.96 (0.94–0.98)	23.8%
	SUV_{AUC}	0.02 ± 0.70	0.95 (0.93–0.98)	27.5%
	SUV_{TBR}	0.76 ± 3.02	0.89 (0.83–0.95)	45.4%
SUV_{mean}	SUV	−0.13 ± 0.76	0.96 (0.94–0.99)	22.3%
	SUV_{AUC}	0.05 ± 0.52	0.95 (0.92–0.98)	27.7%
	SUV_{TBR}	0.68 ± 2.3	0.89 (0.81–0.96)	42.4%

*Mean test–retest differences ± SD.
†Repeatability coefficient for relative differences according to method of Bland and Altman (*27*).
Data in parentheses are 95% CIs.

$P = 0.684$) and location (36.7% [intrathoracic] vs. 34.4% [intraabdominal/pelvic]; $P = 0.820$).

Repeatability of Total Lesion Uptake. The repeatability coefficients for total lesion uptake and area under the total lesion uptake curve were comparable (33% vs. 31%; $P = 0.954$), whereas the total lesion uptake tumor-to-background ratio showed a larger variance of 51% ($P < 0.001$) (Supplemental Fig. 3). No significant difference was found between lesions with a total uptake below or above the median value of 30.9 (repeatability coefficient, 40.9% [<30.9] vs. 23.1% [>30.9]; $P = 0.093$).

The repeatability coefficients for total lesion uptake were independent of uptake (31% [$SUV_{peak} < 8.3$] vs. 34% [$SUV_{peak} > 8.3$]; $P = 0.139$) and MTV (42.1% [<4.2 cm^3] vs. 25.3% [>4.2 cm^3]; $P = 0.037$). The repeatability coefficients for total lesion uptake were also independent of tissue type (30.8% [bone] vs. 35.7% [lymph nodes]; $P = 0.241$) and location (35.5% [intrathoracic] vs. 32.4% [intraabdominal/pelvic]; $P = 0.778$).

An overview of all repeatability coefficients for the semiquantitative parameters (SUV, MTV, and total lesion uptake) as a function of uptake, MTV, and metastatic tissue type and location is presented in Supplemental Tables 4 and 5. Individual-patient repeatability coefficients for all semiquantitative parameters are provided in Supplemental Figures 4–6. The repeatability coefficients were comparable across most subjects. Patient 1 had the poorest repeatability coefficients regarding SUV (~35% for all 9 combinations of SUV types and normalizations). Patients 2, 4, and 7 had the poorest repeatability coefficients regarding MTV (~51% in all 3 patients) and total lesion uptake (55%, 51%, and 67%, respectively).

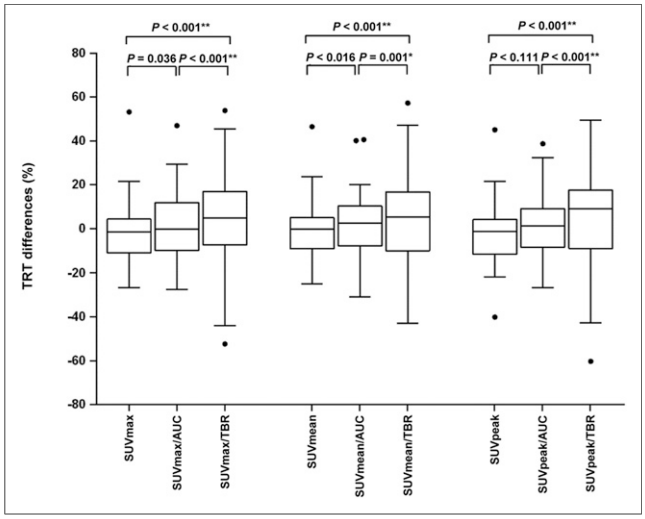

FIGURE 1. Relative differences between test–retest data and mean values for different types of SUVs (SUV_{max}, SUV_{mean}, and SUV_{peak}) and their normalizations (SUV, SUV_{AUC}, and SUV_{TBR}). Bonferroni-corrected significance levels are $P < 0.0056$ (for significance level of 5%, denoted by *) and $P < 0.0011$ (for significance level of 1%, denoted by **). • = outliers (>2 SDs); TRT = test–retest.

DISCUSSION

In a previous study, we investigated ^{18}F-fluoromethylcholine kinetics in metastatic prostate cancer (*24*) and demonstrated that SUV cannot be used to estimate uptake. SUV_{AUC} based on 2 consecutive

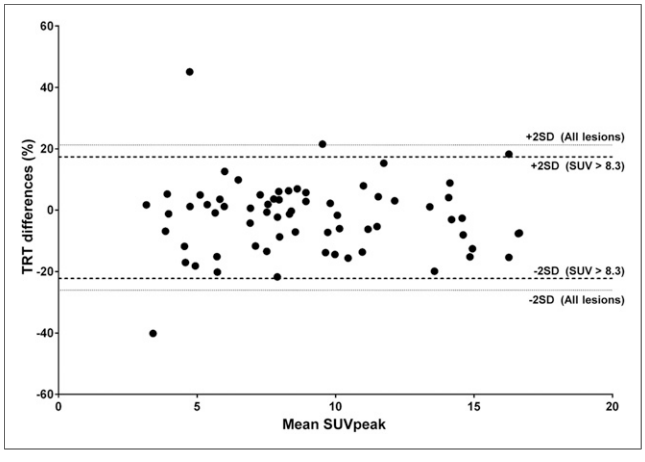

FIGURE 2. Bland–Altman plots for relative differences between test–retest data and mean values for SUV_{peak}.

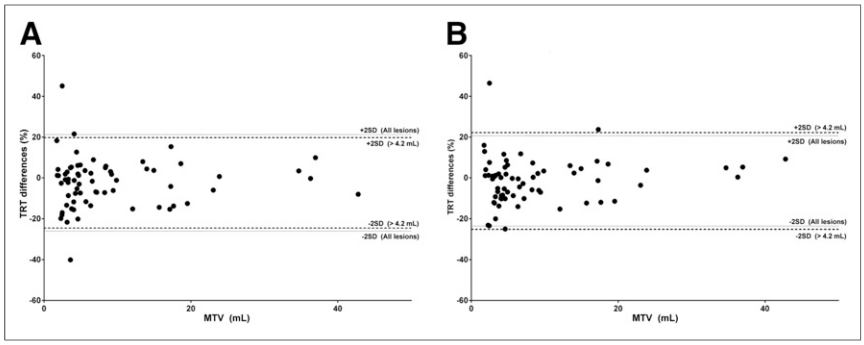

FIGURE 3. Bland–Altman plots for relative test–retest repeatability of SUV_{peak} (A) and SUV_{mean} (B), as function of MTV.

PET scans was proposed as a clinically feasible alternative. In the present study, we prospectively assessed the repeatability of quantitative ^{18}F-fluoromethylcholine PET/CT parameters in patients with prostate cancer and found that the repeatability of SUV_{AUC} is comparable to that of standard SUV and that differences in uptake of 30% or more are likely to represent treatment effects.

Test–retest repeatability is essential for clinical implementation of any parameter of response assessment. Because of the heterogeneous biologic behavior of prostate cancer (*4*), and in light of the rapid evolution of treatment modalities (*14*), biologic markers are needed that adequately monitor response to therapy. The standard treatment-response criteria, RECIST, do not apply to metastatic prostate cancer, further complicating the issue of evaluating response (*30*). A subsequently proposed system for measuring functional response with ^{18}F-FDG PET/CT, PERCIST, might also apply to radiolabeled choline (*31*). Nevertheless, before ^{18}F-fluoromethylcholine PET/CT can be implemented as a biomarker for response evaluation in prostate cancer, the repeatability of the tracer should be known (*32*).

To the best of our knowledge, the repeatability of ^{18}F-fluoromethylcholine measurements in metastatic prostate cancer has not been assessed previously. Pegard et al. (*33*) addressed the reproducibility of observer interpretation (i.e., visual evaluation and classification of focally increased uptake as being malignant or benign) of ^{18}F-fluoromethylcholine PET/CT examinations in patients with biochemically recurrent prostate cancer. The authors found good concordance when evaluating bone metastases and abdominal or pelvic lymphatic recurrences in previously treated patients. A limited usefulness was found at the prostate level in untreated patients.

Observed repeatabilities in our study were within the range seen for other commonly used radiotracers, such as ^{18}F-FDG and ^{18}F-fluorothymidine. In analyzing changes in ^{18}F-FDG uptake, the generally accepted PERCIST response classification assumes a true biologic change when there is a change in SUV_{peak} exceeding 30% in combination with a 0.8-unit change in absolute SUV_{peak} (*31*). In a metaanalysis on the repeatability of ^{18}F-FDG uptake measurements in tumors, de Langen et al. (*34*) identified 8 eligible studies. SUV_{mean} had better repeatability than SUV_{max}. A minimal relative change of 20% in combination with a 1.2-unit change in SUV_{mean} was presumed to represent a biologic change. Comparable results were reported by Rockall et al. (*35*) in a study on the repeatability of quantitative ^{18}F-FDG PET/CT in recurrent ovarian carcinoma. The repeatability coefficients suggested that a decrease from baseline of up to 20% for tumor SUV and up to 15% for tumor size could be used to determine early tumor response. In a study addressing the repeatability and reproducibility of ^{18}F-FDG and ^{18}F-fluorothymidine PET measurements of tumor volume, Hatt et al. (*36*) found comparable percentage differences for the two tracer datasets. Differences larger than 30% were considered indicative of treatment response.

Frings et al. (*26*) analyzed the repeatability of MTV with ^{18}F-FDG and ^{18}F-fluorothymidine in lung cancer. Repeatability was better for larger tumors. For ^{18}F-FDG, changes of more than 37% in lesions larger than 4.2 cm^3 represented a biologic effect. We obtained comparable results when using this MTV threshold to analyze ^{18}F-fluoromethylcholine test–retest data. A larger variability was found in small MTVs ($<$4.2 cm^3), suggesting that a similar lower threshold for MTV should be used in treatment-response studies. In the case of total lesion choline uptake, repeatability coefficients were not significantly different for lesions smaller or larger than 4.2 cm^3 (after correcting for multiple comparisons; Bonferroni-corrected significance level, $P < 0.0056$). However, the uncorrected P value equaled 0.037, suggesting a trend toward poorer repeatability for smaller lesions.

Two recently published papers explored the prognostic value of metabolic parameters when ^{18}F-fluoromethylcholine PET/CT is used in biochemically recurrent prostate cancer (*37*) or castration-resistant prostate cancer (*38*). In a multivariate analysis, Colombié et al. (*37*) identified age less than 70 y, SUV_{mean} greater than or equal to 3, and standardized metabolic activity greater than or equal to 23 as independent prognostic factors for disease-free survival. In a prospective study using ^{18}F-fluoromethylcholine PET/CT, Kwee et al. (*38*) found that whole-body tumor indices based on the quantification of net metabolically active tumor volume and total lesion activity were predictive of overall survival. In our study, we used comparable metrics, with an emphasis on the repeatability of ^{18}F-fluoromethylcholine as a potential biomarker for response evaluation in prostate cancer.

Repeatability coefficients for individual patients were comparable across most subjects. However, patient 1 had the poorest repeatability coefficients regarding SUV, because of a small lesion

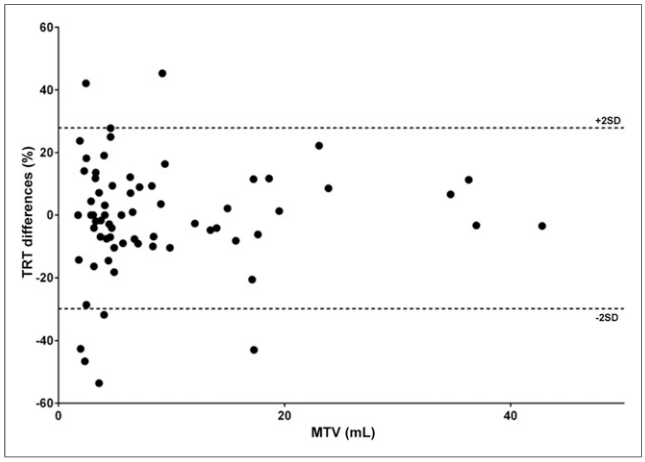

FIGURE 4. Relative differences between test–retest data for MTV as function of MTV.

(1.5-cm short-axis diameter) at the right ventral edge of a vertebral body (Th10), close to the liver. We hypothesize that the relatively large test–retest difference (~35%) of this lesion was caused by incorrect scatter correction due to high image-derived blood activity concentrations near physiologically ^{18}F-fluoromethylcholine–avid structures. Such an effect might lead to large quantification errors with image-derived input function obtained from blood VOI in these areas (*24*). Three patients (2, 4, and 7) had larger repeatability coefficients for MTV and total lesion uptake than the other subjects. This difference is likely explained by difficulty in lesion segmentation, such that errors in MTV are also propagated into poorer repeatability for total lesion uptake. Besides, all these patients presented with small metastatic lesions (~1.5 cm) with slightly increased ^{18}F-fluoromethylcholine uptake.

With respect to the proposed 30% cutoff for SUV measures, the test–retest differences pooled over all lesions or lesions with an SUV of more than 8.3 were approximately 20% (Fig. 2). However, for lesions with an SUV of less than 8.3, the repeatability coefficients were about 30% (Supplemental Table 4). Since, in clinical practice, most patients will have a combination of metastatic lesions with SUVs both above and below 8.3, we decided to adopt the more conservative value of 30%.

A possible limitation of our study is the limited number of subjects. The minimum required sample size was calculated to be 12 patients (with minimally 2 measurements per patient). This sample size yields 80% power for testing the hypothesis ICC $\leq$ 0.6 against the 1-sided alternative (ICC $>$ 0.6) at a significance level of 5% when the true ICC is equal to 0.9 (*39*). Moreover, for this sample size of 12 patients, the confidence intervals for the limits of agreement ranged to approximately 1.1 times the SD of the difference scores, at either side of the estimated limit of agreement (*27*). Thus, the present study provides a reasonable estimate of the expected repeatability coefficients.

CONCLUSION

In patients with metastatic prostate cancer, the repeatability of SUV_{AUC} was comparable to that of standard SUV, indicating that differences in ^{18}F-fluoromethylcholine uptake of 30% or more are likely to represent treatment effects. The repeatability of MTV and total lesion uptake was about 35%. The observed repeatabilities were on the same order of magnitude as those seen for other commonly used radiotracers, such as ^{18}F-FDG and ^{18}F-fluorothymidine.

DISCLOSURE

The costs of publication of this article were defrayed in part by the payment of page charges. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734. No potential conflict of interest relevant to this article was reported.

ACKNOWLEDGMENTS

We thank our colleagues at the Department of Radiology and Nuclear Medicine for assistance with tracer production and data acquisition. We are also grateful to the urologists at VU University Medical Center (André N. Vis, MD, PhD) and Amstelland Hospital Amstelveen (Joop W. Noordzij, MD, and Sven Nadorp, MD), as well as the oncologists at VU University Medical Center (Catharina Menke-van der Houven van Oordt, MD, PhD, Joyce van Dodewaard-de Jong, MD, and Jens Voortman, MD, PhD), for assistance with patient selection and inclusion.

REFERENCES

1. Jemal A, Bray F, Center MM, et al. Global cancer statistics. *CA Cancer J Clin.* 2011;61:69–90.
2. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur J Cancer.* 2013;49:1374–1403.
3. Pienta KJ, Bradley D. Mechanisms underlying the development of androgen-independent prostate cancer. *Clin Cancer Res.* 2006;12:1665–1671.
4. Scher HI, Sawyers CL. Biology of progressive, castration-resistant prostate cancer: directed therapies targeting the androgen-receptor signaling axis. *J Clin Oncol.* 2005;23:8253–8261.
5. Feldman BJ, Feldman D. The development of androgen-independent prostate cancer. *Nat Rev Cancer.* 2001;1:34–45.
6. Tannock IF, De Wit R, Berry WR, et al. Docetaxel plus prednisone or mitoxantrone plus prednisone for advanced prostate cancer. *N Engl J Med.* 2004;351:1502–1512.
7. de Bono JS, Oudard S, Ozguroglu M, et al. Prednisone plus cabazitaxel or mitoxantrone for metastatic castration-resistant prostate cancer progressing after docetaxel treatment: a randomised open-label trial. *Lancet.* 2010;376:1147–1154.
8. Fizazi K, Scher HI, Molina A, et al. Abiraterone acetate for treatment of metastatic castration-resistant prostate cancer: final overall survival analysis of the COU-AA-301 randomised, double-blind, placebo-controlled phase 3 study. *Lancet Oncol.* 2012;13:983–992.
9. Ryan CJ, Smith MR, de Bono JS, et al. Abiraterone in metastatic prostate cancer without previous chemotherapy. *N Engl J Med.* 2013;368:138–148.
10. Scher HI, Fizazi K, Saad F, et al. Increased survival with enzalutamide in prostate cancer after chemotherapy. *N Engl J Med.* 2012;367:1187–1197.
11. Beer TM, Armstrong AJ, Rathkoph DE, et al. Enzalutamide in metastatic prostate cancer before chemotherapy. *N Engl J Med.* 2014;371:424–433.
12. Kantoff PW, Higano CS, Shore ND, et al. Sipuleucel-T immunotherapy for castration-resistant prostate cancer. *N Engl J Med.* 2010;363:411–422.
13. Parker C, Nilsson S, Heinrich D, et al. Alpha emitter radium-223 and survival in metastatic prostate cancer. *N Engl J Med.* 2013;369:213–223.
14. van Dodewaard-de Jong JM, Verheul HM, Bloemendal HJ, de Klerk JM, Carducci MA, van den Eertwegh AJ. New treatment options for patients with metastatic prostate cancer: what is the optimal sequence? *Clin Genitourin Cancer.* 2015;13:271–279.
15. Bray F, Lortet-Tieulent J, Ferlay J, et al. Prostate cancer incidence and mortality trends in 37 European countries: an overview. *Eur J Cancer.* 2010;46:3040–3052.
16. Heidenreich A, Bastian PJ, Bellmunt J, et al. EAU guidelines on prostate cancer. Part 1: screening, diagnosis, and local treatment with curative intent—update 2013. *Eur Urol.* 2014;65:124–137.
17. Schiepers C, Dahlbom M. Molecular imaging in oncology: the acceptance of PET/CT and the emergence of MR/PET imaging. *Eur Radiol.* 2011;21:548–554.
18. Bauman G, Belhocine T, Kovacs M, et al. ^{18}F-fluorocholine for prostate cancer imaging: a systematic review of the literature. *Prostate Cancer Prostatic Dis.* 2012;15:45–55.
19. Fuccio C, Rubello D, Castellucci P, et al. Choline PET/CT for prostate cancer: main clinical applications. *Eur J Radiol.* 2011;80:e50–e56.
20. Umbehr MH, Müntener M, Hany T, et al. The role of ^{11}C-choline and ^{18}F-fluorocholine positron emission tomography (PET) and PET/CT in prostate cancer: a systematic review and meta-analysis. *Eur Urol.* 2013;64:106–117.
21. Evangelista L, Zattoni F, Guttilla A, et al. Choline PET or PET/CT and biochemical relapse of prostate cancer: a systematic review and meta-analysis. *Clin Nucl Med.* 2013;38:305–314.
22. Müller SA, Holzapfel K, Seidl C, et al. Characterization of choline uptake in prostate cancer cells following bicalutamide and docetaxel treatment. *Eur J Nucl Med Mol Imaging.* 2009;36:1434–1442.
23. Oprea-Lager DE, van Kanten MP, van Moorselaar RJ, et al. [^{18}F]fluoromethylcholine as a chemotherapy response read-out in prostate cancer cells. *Mol Imaging Biol.* 2015;17:319–327.
24. Verwer EE, Oprea-Lager DE, van den Eertwegh AJ, et al. Quantification of ^{18}F-fluorocholine kinetics in patients with prostate cancer. *J Nucl Med.* 2015;56:365–371.
25. Boellaard R, Delgado-Bolton R, Oyen WJ, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging—version 2.0. *Eur J Nucl Med Mol Imaging.* 2015;42:328–354.
26. Frings V, de Langen AJ, Smit EF, et al. Repeatability of metabolically active volume measurements with ^{18}F-FDG and ^{18}F-FLT PET in non-small cell lung cancer. *J Nucl Med.* 2010;51:1870–1877.

27. Bland JM, Altman DG. Agreement between methods of measurement with multiple observations per individual. *J Biopharm Stat.* 2007;17:571–582.
28. García-Pérez MA. Statistical criteria for parallel tests: a comparison of accuracy and power. *Behav Res Methods.* 2013;45:999–1010.
29. Keselman HJ, Wilcox RR, Algina J, Othman AR, Fradette K. A comparative study of robust tests for spread: asymmetric trimming strategies. *Br J Math Stat Psychol.* 2008;61:235–253.
30. Wallace TJ, Torre T, Grob M, et al. Current approaches, challenges and future directions for monitoring treatment response in prostate cancer. *J Cancer.* 2014;5: 3–24.
31. 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–150S.
32. Morisson C, Jerai R, Liu G. Imaging of castration-resistant prostate cancer: development of imaging response biomarkers. *Curr Opin Urol.* 2013;23:230–236.
33. Pegard C, Gallazzini-Crépin C, Giai J, et al. Study of inter- and intra-observer reproducibility in the interpretation of [^{18}F]choline PET/CT examinations in patients suffering from biochemically recurrent prostate cancer following curative treatment. *EJNMMI Res.* 2014;4:25.
34. de Langen AJ, Vincent A, Velasquez LM, et al. Repeatability of ^{18}F-FDG uptake measurements in tumors: a meta-analysis. *J Nucl Med.* 2012;53:701–708.
35. Rockall AG, Avril N, Lam R, 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;20:2751–2760.
36. Hatt M, Cheze-Le Rest C, Aboagye EO, et al. Reproducibility of ^{18}F-FDG and 3′-deoxy-3′-^{18}F-fluorothymidine PET tumor volume measurements. *J Nucl Med.* 2010;51:1368–1376.
37. Colombié M, Campion L, Bailly C, et al. Prognostic value of metabolic parameters and clinical impact of ^{18}F-fluorocholine PET/CT in biochemical recurrent prostate cancer. *Eur J Nucl Med Mol Imaging.* 2015;42:1784–1793.
38. Kwee SA, Lim J, Watanabe A, Kromer-Baker K, Coel MN. Prognosis related to metastatic burden measured by ^{18}F-fluorocholine PET/CT in castration-resistant prostate cancer. *J Nucl Med.* 2014;55:905–910.
39. Shoukri MM, Asyali MH, Donner A. Sample size requirements for the design of reliability study: a review and new results. *Stat Methods Med Res.* 2004;13:251–271.