Repeatability of ^{18}F-FLT PET in a Multicenter Study of Patients with High-Grade Glioma

Martin A. Lodge[1], Matthias Holdhoff[2], Jeffrey P. Leal[1], Asim K. Bag[3], L. Burt Nabors[3], Akiva Mintz[4], Glenn J. Lesser[4], David A. Mankoff[5], Arati S. Desai[5], James M. Mountz[6], Frank S. Lieberman[6], Joy D. Fisher[2], Serena Desideri[2], Xiaobu Ye[2], Stuart A. Grossman[2], David Schiff[7], and Richard L. Wahl[1,2]

[1]The Russell H. Morgan Department of Radiology and Radiological Science, Johns Hopkins University School of Medicine, Baltimore, Maryland; [2]Brain Cancer Program, Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins, Baltimore, Maryland; [3]University of Alabama, Birmingham, Alabama; [4]Wake Forest University School of Medicine, Winston Salem, North Carolina; [5]University of Pennsylvania, Philadelphia, Pennsylvania; [6]University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania; and [7]University of Virginia, Charlottesville, Virginia

Quantitative 3′-deoxy-3′-^{18}F-fluorothymidine (^{18}F-FLT) PET has potential as a noninvasive tumor biomarker for the objective assessment of response to treatment. To guide interpretation of these quantitative data, we evaluated the repeatability of ^{18}F-FLT PET as part of a multicenter trial involving patients with high-grade glioma. **Methods:** ^{18}F-FLT PET was performed on 10 patients with recurrent high-grade glioma at 5 different institutions within the Adult Brain Tumor Consortium trial ABTC1101. Data were acquired according to a double baseline protocol in which PET examinations were repeated within 2 d of each other with no intervening treatment. On each of the 2 imaging days, dedicated brain PET was performed at 2 time points, 1 and 3 h after ^{18}F-FLT administration. Tumor SUVs and related parameters were measured at a central laboratory using various volumes of interest: isocontour at 30% of the maximum pixel ($SUV_{mean\_30\%}$), gradient-based segmentation ($SUV_{mean\_gradient}$), the maximum pixel (SUV_{max}), and a 1-mL sphere at the region of highest uptake (SUV_{peak}). Repeatability coefficients (RCs) were calculated from the relative differences between corresponding SUV measurements obtained on the 2 d. **Results:** RCs for tumor SUVs were 22.5% ($SUV_{mean\_30\%}$), 23.8% ($SUV_{mean\_gradient}$), 23.2% (SUV_{max}), and 18.5% (SUV_{peak}) at 1 h after injection. Corresponding data at 3 h were 22.4%, 25.0%, 27.3%, and 23.6%. Normalizing the tumor SUV data with reference to a background region improved repeatability, and the most stable parameter was the tumor-to-background ratio derived using SUV_{peak} (RC, 16.5%). **Conclusion:** SUV quantification of ^{18}F-FLT uptake in glioma had an RC in the range of 18%–24% when imaging began 1 h after ^{18}F-FLT administration. The volume-of-interest methodology had a small but not negligible influence on repeatability, with the best performance obtained using SUV_{peak}. Although changes in ^{18}F-FLT SUV after treatment cannot be directly interpreted as a change in tumor proliferation, we have established ranges beyond which SUV differences are likely due to legitimate biologic effects.

Key Words: repeatability; reproducibility; FLT; PET; SUV; glioma

J Nucl Med 2017; 58:393–398
DOI: 10.2967/jnumed.116.178434

The radiolabeled thymidine analog 3′-deoxy-3′-^{18}F-fluorothymidine (^{18}F-FLT) is incorporated into cells and phosphorylated by thymidine kinase 1. Its retention within cells reflects, in part, thymidine kinase activity and is often positively correlated with cellular proliferation. ^{18}F-FLT PET can allow noninvasive assessment of tumor proliferation (*1*) and has been used for monitoring response to treatment in various malignancies (*2–5*). The tracer would seem to be well suited to the study of glioma as it is associated with low uptake in normal brain, frequently leading to high tumor-to-background (T-to-B) contrast compared with the glucose analog ^{18}F-FDG. However, the use of ^{18}F-FLT for response assessment in gliomas is complicated because its uptake can be influenced by multiple mechanisms.

^{18}F-FLT does not readily cross the intact blood–brain barrier (BBB), but uptake is relatively high in brain tumors with a disrupted BBB (*6,7*). The distribution of ^{18}F-FLT closely matches the region of contrast enhancement on postgadolinium T1-weighted MRI (*8*), indicating that ^{18}F-FLT uptake is strongly influenced by BBB permeability. Sustained retention of ^{18}F-FLT in tumors long after gadolinium-based contrast agents have washed out suggests an additional trapping process. Several papers (*6,9–12*) report a correlation between ^{18}F-FLT uptake in brain tumors, measured using SUV, and the histopathologic proliferation marker Ki-67. The correlation is usually quite weak, consistent with the thought that ^{18}F-FLT uptake is driven not just by proliferation but by additional mechanisms, such as BBB permeability and tumor blood flow. The uptake of ^{18}F-FLT in brain tumors is thus a complex process, and interpretation of the data is not straightforward.

Because of the multiple mechanisms that influence ^{18}F-FLT uptake in glioma, careful consideration has been given to the optimum quantitative analysis approach. Tracer kinetic modeling has the potential to distinguish transport effects from proliferation, and encouraging results have been presented (*13*). A 3-compartment model has been adopted (Supplemental Fig. 1; supplemental materials are available at http://jnm.snmjournals.org), and both k_3 and the overall influx constant (K_i) have been shown to be correlated with the in vitro proliferation index Ki-67 (*14*). However, with up to 5 unknown model parameters, including 4 rate constants and a blood volume term, estimates of k_3 are often unreliable (*15–17*). K_i is more robust but this parameter reflects the overall uptake rate and includes contributions due to blood flow and BBB permeability. As such, changes in K_i do not necessarily indicate altered tumor cell

Received May 17, 2016; revision accepted Aug. 28, 2016.
For correspondence or reprints contact: Martin A. Lodge, Division of Nuclear Medicine and Molecular Imaging, 601 North Caroline St., Baltimore, MD 21287.
E-mail: mlodge1@jhmi.edu
Published online Sep. 29, 2016.

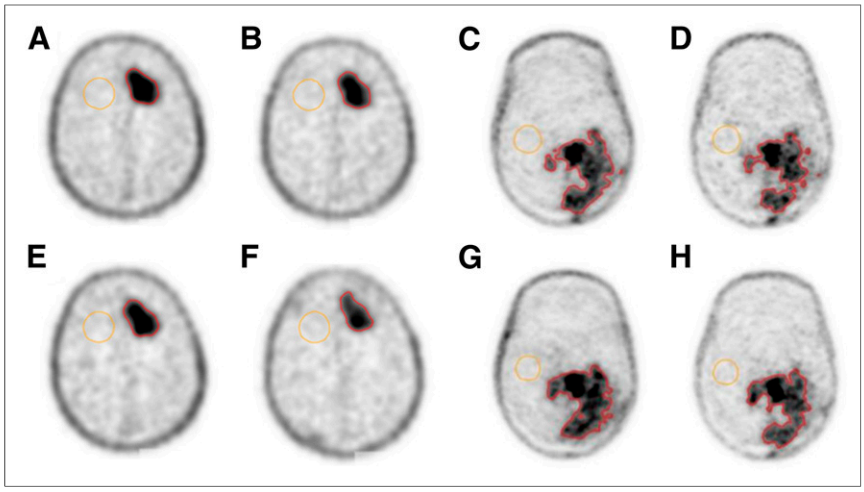

FIGURE 1. Example ^{18}F-FLT PET transaxial images for 2 different patients, acquired at 4 different time points: day 1, 1 h (A and C) and 3 h (B and D); day 2, 1 h (E and G) and 3 h (F and H). Images are shown in a common (inverse) gray scale. Isocontour tumor VOIs and normal background VOIs are shown in color.

proliferation because potential changes in BBB disruption or blood flow may also contribute. Simple uptake measures such as SUV are highly correlated with K_i (*18*) and have been proposed for therapy monitoring (*19*). Despite not being capable of dissecting the complexity of ^{18}F-FLT kinetics, SUV has been used successfully for predicting response to treatment in patients with high-grade glioma (*20,21*). Because SUV does not distinguish between proliferation and permeability effects, the observed changes in the ^{18}F-FLT signal with treatment may have been driven, at least in part, by the permeability alterations that were seen on MRI. Nevertheless, changes in ^{18}F-FLT SUV after treatment with bevacizumab were highly predictive of progression-free and overall survival and appear to be slightly more predictive than anatomic MRI for early assessment (*21*).

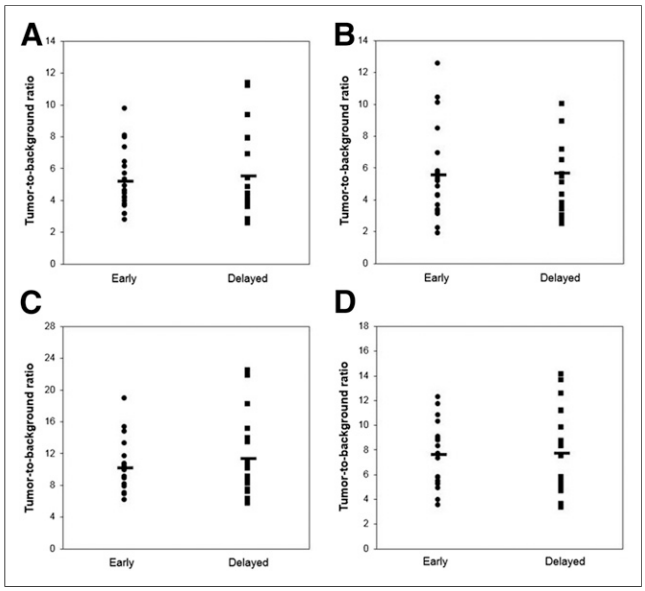

FIGURE 2. T-to-B ratio for 4 different tumor SUVs: $SUV_{mean\_30\%}$ (A), $SUV_{mean\_gradient}$ (B), SUV_{max} (C), and SUV_{peak} (D). Labels Early and Delayed refer to data acquired at approximately 1 and 3 h after injection. Horizontal markers indicate mean T-to-B ratios.

Further studies are required to clarify the role of ^{18}F-FLT in glioma response assessment, but initial reports suggest a performance similar to the current MRI standard based on the Response Assessment in Neuro-Oncology Working Group criteria (*22*). These criteria involve quantitative measurements of tumor size, but also incorporate a qualitative evaluation of T2 fluid-attenuated inversion recovery images, presence or absence of new lesions on MRI, and clinical factors. In a similar manner, future ^{18}F-FLT PET response criteria could combine image-derived quantitative metrics with clinical factors such as corticosteroid dose and clinical status. Large T-to-B ratios mean that quantitative analysis could potentially be highly automated, objective, and reproducible, making ^{18}F-FLT PET well suited for multicenter trials. Successful multicenter deployment requires standardization of methodology and an awareness of the repeatability of the technique (*23*). Protocols that involve tracer kinetic modeling are difficult to implement in a multicenter setting because they require dynamic image acquisition in conjunction with blood sampling, radioactive sample measurement, and metabolite correction. Although associated with the limitations noted previously, static PET acquisition and SUV quantification currently provide the most practical approach for multicenter trials. However, several different variants of SUV exist, and the optimum methodology for response assessment is not clear. For example, the previously mentioned study of early outcome prediction (*21*) used an SUV that was somewhat similar to SUV_{peak}, as opposed to the more common SUV_{max}. The present study measures the repeatability of various ^{18}F-FLT PET metrics, using data acquired within a multicenter consortium. In particular, we focus on methodologic issues and their implications for response assessment.

MATERIALS AND METHODS

Study Design

The repeatability of quantitative ^{18}F-FLT PET was assessed using a prospective, multicenter, test–retest study design. Ten patient volunteers (9 men, 1 woman) with recurrent high-grade glioma participated (World Health Organization grade IV, $n = 6$; grade III, $n = 4$). All patient volunteers showed evidence of tumor enhancement on postgadolinium T1-weighted MRI and had no treatment for 3 mo before study commencement. Individual patients underwent 2 ^{18}F-FLT PET studies, each performed on consecutive days with no intervening treatment. These data were acquired over 5 different centers as part of an Adult Brain Tumor Consortium trial (ABTC1101, ClinicalTrials.gov identifier NCT01480050) that was approved by the Institutional Review Boards of the various institutions. All patients provided written informed consent to participate in the study.

Data Acquisition

The study included data acquired on 5 different commercial PET/CT systems: Ingenuity TF (Philips) and Discovery 710, Discovery ST, Discovery STE, and Discovery VCT (GE Healthcare). Quality assurance images from each system were acquired before patient imaging using an ^{18}F-filled 20-cm-diameter cylindric water phantom. All scanners produced phantom images with average SUVs in the range 0.95–1.05 g/mL and had no artifacts or nonuniformities on visual inspection. Although

different scanners were involved, individual patients were studied using the same scanner system on each of their 2 imaging days. Data acquisition on each day proceeded according to an identical imaging protocol. ^{18}F-FLT (2.6 MBq/kg) was administered, and PET/CT data were acquired at 2 different time points, 1 and 3 h after injection. The test–retest comparison was between corresponding time points acquired on consecutive days, not between the images acquired 1 and 3 h after injection. The 2 time points allowed parallel measurements of repeatability at 2 different, clinically relevant levels of image statistical quality.

At both the 1- and the 3-h time points, the patient's head was positioned in the center of the field of view and carefully supported using the scanner manufacturer's standard head-holder. Immediately before each PET acquisition, low-dose, noncontrast, nondiagnostic CT was acquired for attenuation correction according to the local procedures of the participating sites. PET data were acquired for 10 min as a static scan in 3-dimensional mode. Images were reconstructed using iterative reconstruction in conjunction with corrections for attenuation, scatter, randoms, and detector normalization. Because of the different scanner models and available software options, the reconstruction parameters were not identical among sites but were consistently used for a given scanner. Reconstructed voxel sizes were typically 2 mm in transverse planes, 3.3 mm in the axial direction, and were similar among all scanners.

Image Analysis

Digital images from all sites were transferred to a central laboratory and analyzed using an identical analysis protocol. Initial quality evaluation included checks of manually entered SUV information (activities, times, weights) and visual assessment of image quality including motion artifacts. Quantitative SUV analysis (body weight normalization) used volumes of interest (VOIs) defined in tumor and normal brain. Tumor VOIs were determined using isocontour segmentation (30% of the maximum tumor voxel) and a gradient edge detection approach (*24*). The isocontour VOIs were determined automatically after a spheric guiding volume was manually specified. The gradient edge detection approach involved operator definition of the starting points for gradient segmentation. The activity concentration from all voxels within these VOIs were averaged to obtain $SUV_{mean\_30\%}$ and $SUV_{mean\_gradient}$, respectively. Additional SUV metrics were derived from the maximum voxel within the tumor (SUV_{max}) and from a 1-mL spheric VOI (SUV_{peak}). SUV_{peak} was measured using the XD3 software (Mirada Medical), which automatically positioned a 1-mL sphere within the tumor so as to maximize its average value (*25*). Background SUV was determined by manually placing a 3-cm-diameter sphere in a normal brain region, approximately contralateral to the tumor site and recording the mean value (SUV_{normal}). T-to-B ratios, tumor volume, and the product of SUV_{mean} and volume (akin to total lesion glycolysis for ^{18}F-FDG) were also computed. All VOIs were recalculated for each image volume, as opposed to copying a previously defined VOI from a prior study.

Statistical Analysis

The primary analysis involved a measurement of the repeatability of tumor SUV measurements, although non-SUV metrics and normal brain SUV were also assessed. Each metric (e.g., $SUV_{mean\_30\%}$, $SUV_{mean\_gradient}$, SUV_{max}, SUV_{peak}) at both postinjection scan times (1 and 3 h) were analyzed as follows. For each patient, the SUV difference d was given simply by,

$$d = SUV_2 - SUV_1 \qquad \text{Eq. 1}$$

and the average SUV by,

$$\mu = \frac{1}{2}(SUV_1 + SUV_2). \qquad \text{Eq. 2}$$

SUV_1 and SUV_2 were corresponding SUV measurements of the same tumor made under test–retest conditions, for example, SUV_{max} on day 1 at 1 h after injection and SUV_{max} on day 2 at 1 h after injection. The magnitude of d was found to be proportional to μ for all SUV metrics so the difference data were expressed in relative units as follows,

$$D = \frac{SUV_2 - SUV_1}{\mu} \times 100. \qquad \text{Eq. 3}$$

These data were presented as a Bland–Altman plot (*26*) in which the relative difference (D) was plotted as a function of the SUV average (μ). The within-subject coefficient of variation (wCV) was given by DSD/$\sqrt{2}$, where DSD is the SD of D. Note that wCV reflects the variability in a single measurement, and the $1/\sqrt{2}$ factor accounts for the fact that D is subject to noise in both SUV_1 and SUV_2. The repeatability coefficient (RC) was given by 1.96 × DSD (*26*) under the assumption that D was normally distributed (Shapiro–Wilk test). RC represents the 95% limits of repeatability for the difference between 2 SUV measurements made under test–retest conditions. In other words, for a perfectly stable tumor, the relative difference between the SUV measurements derived from 2 separate PET/CT studies should be expected to be within ± RC, 95% of the time.

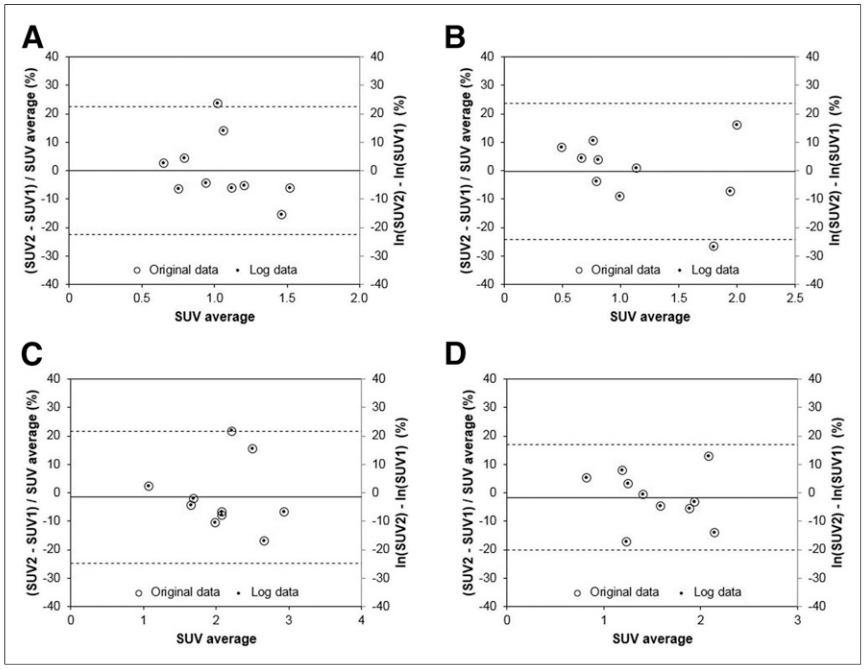

FIGURE 3. Bland–Altman plots for SUV data acquired at 1 h after injection. Data for 4 different VOI schemes are shown: $SUV_{mean\_30\%}$ (A), $SUV_{mean\_gradient}$ (B), SUV_{max} (C), and SUV_{peak} (D). Note data are plotted on both primary (left) and secondary (right) *y*-axes as open circles and smaller solid circles, respectively. Dashed lines indicate 95% limits of repeatability. Quantity μ was expressed as SUV average for clarity.

RESULTS

All 10 patients completed the 2-d imaging protocol with no missing data or

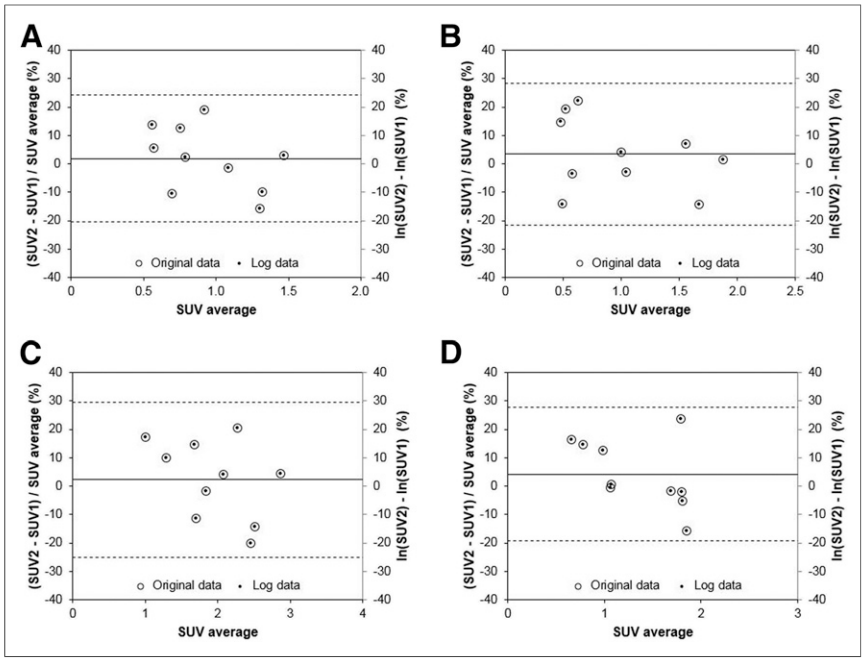

FIGURE 4. Bland–Altman plots for SUV data acquired at 3 h after injection. Data for 4 different VOI schemes are shown: $SUV_{mean\_30\%}$ (A), $SUV_{mean\_gradient}$ (B), SUV_{max} (C), and SUV_{peak} (D). The data are plotted on both primary (left) and secondary (right) *y*-axes as open circles and smaller solid circles, respectively. Dashed lines indicate 95% limits of repeatability. Quantity μ was expressed as SUV average for clarity.

incomplete datasets. No data were rejected due to poor image quality, motion problems, or other reasons. The mean injected activity was 208 ± 41 MBq. The mean postinjection scan times were 62 ± 3 and 175 ± 17 min on day 1 and 61 ± 2 and 171 ± 12 min on day 2. Figure 1 shows example ^{18}F-FLT PET images from 2 patients with tumor and normal brain VOIs overlaid in color. No manual adjustments were needed for the isocontour VOIs, although the gradient-based VOIs required some manual editing to correct clearly erroneous segmentation in 4 of the 10 patients. Figure 2 shows the T-to-B ratios (SUV_x/SUV_{normal}) for each tumor SUV metric at each of the 2 postinjection scan times. The T-to-B ratios were similar at both time points (x = mean_30%, 5.2 vs. 5.5; x = mean_gradient, 5.5 vs. 5.7; x = peak, 7.6 vs. 7.7, paired *t* tests $P > > 0.05$), although there was a statistically significant increase at the later time for SUV_{max} (10.2 vs. 11.3, paired *t* test $P = 0.04$).

Figures 3 and 4 show Bland–Altman plots for the tumor SUV data acquired at 1 and 3 h, respectively. The difference data D were consistent with the normal distribution for all SUV metrics (Shapiro–Wilk test $P > > 0.05$). Note that secondary *y*-axes were added on the right side of each graph. On these secondary axes, the SUV difference data were plotted after natural log transformation. The mean difference between the log-transformed data and the original data (D) was only 0.1% for SUV_{max} at 1 h. Table 1 shows wCV and RC data for the various SUV parameters at the 2 different time points. Table 2 shows all tumor metrics but in this case the 1- and 3-h data were combined to minimize the impact of outliers in these sparse datasets. Note that no individual data points were excluded from this analysis. Parameters derived from SUV_{peak} had slightly better repeatability than the other tumor sampling methods: SUV_{peak} at 1 h after injection had an RC of 18.5%; SUV_{peak}/SUV_{normal} had an RC of 16.5%.

TABLE 1
Repeatability Metrics for Various SUV Parameters

Parameter	wCV (%) 1 h	RC (%) 1 h	wCV (%) 3 h	RC (%) 3 h
$SUV_{mean\_30\%}$	8.1	22.5	8.1	22.4
$SUV_{mean\_gradient}$	8.6	23.8	9.0	25.0
SUV_{max}	8.4	23.2	9.9	27.3
SUV_{peak}	6.7	18.5	8.5	23.6
SUV_{normal}	5.0	14.0	6.7	18.4

DISCUSSION

In this project, we measured the repeatability of SUV quantification in a multicenter ^{18}F-FLT PET study involving patients with high-grade glioma. Various VOI techniques were evaluated with respect to their repeatability, and although each SUV metric was broadly similar, some interesting trends were observed. At the 1-h time point, the repeatability of the 2 SUV_{mean} parameters were similar to SUV_{max}, despite involving very different volumetric sampling. The greater volume averaging associated with SUV_{mean} may have been offset by greater variability in VOI definition, particularly for gradient-based segmentation. At the later time point, after an additional half-life of radioactive decay, lower counts led to poorer repeatability for SUV_{max}. In contrast, SUV_{mean} averaged a much greater number of voxels and repeatability was similar for the higher count (1 h) and lower count (3 h) images. Of the tumor SUV parameters, the best repeatability was observed for SUV_{peak} at the 1-h time point, probably because it incorporates a degree of volumetric averaging but does not require accurate tumor delineation. The RC for SUV_{peak} increased at the later time point, likely due to greater statistical noise in the lower count images, although SUV_{peak} repeatability remained superior to that of SUV_{max}. Normalizing tumor SUV data with reference to a background region improved repeatability, and the most stable parameter was the T-to-B ratio derived using SUV_{peak}.

Although the repeatability values shown in Tables 1 and 2 correspond to the current ^{18}F-FLT application, the relative trends between SUV parameters may be generally applicable to other imaging situations. The advantage of SUV_{peak} over SUV_{max}, albeit a modest advantage, is consistent with previous studies involving ^{18}F-FDG (*27,28*). It should be pointed out that although this advantage might be expected, given the greater volume averaging associated with SUV_{peak}, other ^{18}F-FDG studies have found that the repeatability of SUV_{max} and SUV_{peak} was broadly comparable (*29,30*). These data were acquired in large multicenter studies, and it is possible that other sources of variability may

TABLE 2
Repeatability Metrics for Various SUV and Other Tumor Parameters

Parameter	wCV (%)	RC (%)
$SUV_{mean\_30\%}$	7.9	21.9
$SUV_{mean\_gradient}$	8.7	24.0
SUV_{max}	9.0	25.0
SUV_{peak}	7.7	21.4
T-to-B ratio ($SUV_{mean\_30\%}/SUV_{normal}$)	7.3	20.1
T-to-B ratio ($SUV_{mean\_gradient}/SUV_{normal}$)	8.3	23.1
T-to-B ratio (SUV_{max}/SUV_{normal})	8.1	22.6
T-to-B ratio (SUV_{peak}/SUV_{normal})	6.0	16.5
Volume (30% isocontour)	11.5 *	32.0 *
Volume (gradient)	28.8 *	79.9 *
$SUV_{mean\_30\%}$ × volume (30% isocontour)	8.2	22.6
$SUV_{mean\_gradient}$ × volume (gradient)	28.9 *	80.1 *

*Data that were not consistent with the normal distribution (Shapiro–Wilk test).
For each tumor parameter, difference data D obtained at 1 and 3 h after injection were combined and analyzed as single dataset to increase statistical power.

have dominated. It is interesting to note that the repeatability of ^{18}F-FLT seems to be much better than corresponding values for ^{18}F-FDG. Whereas the RC for ^{18}F-FLT SUV_{max} (1 h after injection) is around 23%, similar values for ^{18}F-FDG are around 33%, for example, wCV 11.9% (Table 5 in Velasquez et al. (*29*)) and DSD 17% (Table 2 in Weber et al. (*30*)). In our study, we found very little changes in ^{18}F-FLT uptake over the 1- to 3-h time period, and this stability may have contributed to the improved repeatability compared with ^{18}F-FDG, which is known to change substantially over time (*31*).

The repeatability measured in this study (RC = 23.2% for SUV_{max}, 1 h after injection) was consistent with previous publications involving ^{18}F-FLT. De Langen et al. (*32*) reported a DSD of 11% for SUV_{max}, corresponding to an RC of 21.6% (non–small cell lung cancer, $n = 9$; head and neck cancer, $n = 6$). Hatt et al. (*33*) reported an RC for SUV_{max} of 29.2% in breast cancer ($n = 9$). Kenny et al. (*4*) reported a DSD of 10.5% for SUV_{mean} in breast cancer ($n = 8$), which corresponds to an RC of 20.6%. Shields et al. (*34*) used a repeatability metric that is not directly comparable to RC. However, they measured relative differences in SUV_{mean} of up to 21% (non–small cell lung cancer, $n = 9$). The current study may be the first to evaluate ^{18}F-FLT repeatability in a multicenter setting, and it is encouraging that our results are so similar to previous single-center findings. Also the repeatability measured in the present study supports the 25% threshold for SUV change that has been used previously (*21*). An obvious limitation of our work is the small number of patients ($n = 10$) who were able to participate, although this limitation is not uncommon in studies of this sort.

As a point of interest, we show a secondary scale in Figures 3 and 4 in which the test–retest data were plotted after log transformation. Log transformation of this kind is often used when the magnitude of the difference data is proportional to the mean (*26*). Unfortunately, presentation of data on the log scale can seem confusing and unintuitive. The basic properties of log transformation indicate that the difference of log data is related to the log of the quotient as follows:

$$\ln(SUV_2) - \ln(SUV_1) = \ln\left(\frac{SUV_2}{SUV_1}\right). \quad \text{Eq. 4}$$

Yet despite this simple relationship, it is not obvious how to interpret the difference of 2 SUVs after log transformation. In Figures 3 and 4, we show the data plotted according to Equation 3 and also according to Equation 4 (×100 for interpretation as a percentage). It can be seen that the difference between 2 SUVs after natural log transformation (small filled circles) is almost exactly equivalent to the relative difference defined by Equation 3 (open circles). For example, if SUV_1 and SUV_2 are assumed to be 4 and 5, respectively $(SUV_2 - SUV_1)$/average $(SUV_1, SUV_2) = 0.222$ and $\ln(SUV_2) - \ln(SUV_1) = 0.223$. The difference between natural logs can thus be directly interpreted without the need for back-transformation (*35*). So rather than being unintuitive, the difference between natural logs actually has a natural interpretation.

CONCLUSION

SUV quantification of ^{18}F-FLT uptake in glioma has an RC in the range of 18%–24% when imaging began 1 h after ^{18}F-FLT administration. The VOI methodology had a small but not negligible influence on repeatability, with the best performance obtained using SUV_{peak}. Although changes in ^{18}F-FLT SUV after treatment cannot be directly interpreted as a change in tumor proliferation, we have established ranges beyond which SUV differences are likely due to legitimate biologic effects.

DISCLOSURE

This study was supported by the NCI-funded Adult Brain Tumor Consortium grant UM1CA137443, Sidney Kimmel Comprehensive Cancer Center core grant 5P30CA006973-52, NCI grant 5U01CA140204-05, and a research grant from Cavion LLC. The ClinicalTrials.gov identifier is NCT01480050. No other potential conflict of interest relevant to this article was reported.

ACKNOWLEDGMENTS

We thank the patient volunteers for their essential contributions, and we acknowledge the substantial roles of the staff at the participating centers.

REFERENCES

1. Shields AF, Grierson JR, Dohmen BM, et al. Imaging proliferation in vivo with [F-18]FLT and positron emission tomography. *Nat Med.* 1998;4:1334–1336.
2. Herrmann K, Wieder HA, Buck AK, et al. Early response assessment using 3′-deoxy-3′-[^{18}F]fluorothymidine positron emission tomography in high-grade non-Hodgkins lymphoma. *Clin Cancer Res.* 2007;13:3552–3558.

3. Troost EG, Bussink J, Hoffmann AL, Boerman OC, Oyen WJ, Kanders HA. ^{18}F-FLT PET/CT for early response monitoring and dose escalation in oropharyngeal tumors. *J Nucl Med.* 2010;51:866–874.
4. Kenny L, Coombes RC, Vigushin DM, Al-Nahhas A, Shousha S, Aboagye EO. Imaging early changes in proliferation at 1 week post chemotherapy: a pilot study in breast cancer patients with 3′-deoxy-3′-[^{18}F]fluorothymidine positron emission tomography. *Eur J Nucl Med Mol Imaging.* 2007;34:1339–1347.
5. Sohn HJ, Yang YJ, Ryu JS, et al. [^{18}F]fluorothymidine positron emission tomography before and 7 days after Gefitinib treatment predicts response in patients with advanced adenocarcinoma of the lung. *Clin Cancer Res.* 2008;14:7423–7429.
6. Saga T, Kawashima H, Araki N, et al. Evaluation of primary brain tumors with FLT-PET: usefulness and limitations. *Clin Nucl Med.* 2006;31:774–780.
7. Hong IK, Kim JH, Ra YS, Kwon DH, Oh SJ, Kim JS. Diagnostic usefulness of 3′-deoxy-3′-[^{18}F]fluorothymidine positron emission tomography in recurrent brain tumor. *J Comput Assist Tomogr.* 2011;35:679–684.
8. Weber MA, Henze M, Tuttenberg J, et al. Biopsy targeting gliomas: do functional imaging techniques identify similar target areas? *Invest Radiol.* 2010;45:755–768.
9. Chen W, Cloughesy T, Kamdar N, et al. Imaging proliferation in brain tumors with ^{18}F-FLT PET: comparison with ^{18}F-FDG. *J Nucl Med.* 2005;46:945–952.
10. Hatakeyama T, Kawai N, Nishiyama Y, et al. ^{11}C-methionine (MET) and ^{18}F-fluorothymidine (FLT) PET in patients with newly diagnosed glioma. *Eur J Nucl Med Mol Imaging.* 2008;35:2009–2017.
11. Choi SJ, Kim JS, Kim JH, et al. [^{18}F]3′-deoxy-3′-fluorothymidine PET for the diagnosis and grading of brain tumors. *Eur J Nucl Med Mol Imaging.* 2005;32:653–659.
12. Toyota Y, Miyake K, Kawai N, et al. Comparison of 4′-[methyl-^{11}C]thiothymidine (^{11}C-4DST) and 3′-deoxy-3′-[^{18}F]fluorothymidine (^{18}F-FLT) PET/CT in human brain glioma imaging. *EJNMMI Res.* 2015;5:1–8.
13. Spence AM, Muzi M, Link JM, et al. NCI-sponsored trial for the evaluation of safety and preliminary efficacy of 3′-deoxy-3′-[^{18}F]fluorothymidine (FLT) as a marker of proliferation in patients with recurrent gliomas: preliminary efficacy studies. *Mol Imaging Biol.* 2009;11:343–355.
14. Ullrich R, Backes H, Li H, et al. Glioma proliferation as assessed by 3′-fluoro-3′-deoxy-L-thymidine positron emission tomography in patients with newly diagnosed high-grade glioma. 2008;14:2049-2055.
15. Jacobs AH, Thomas A, Kracht LW, et al. ^{18}F-fluoro-L-thymidine and ^{11}C-methylmethionine as markers of increased transport and proliferation in brain tumors. *J Nucl Med.* 2005;46:1948–1958.
16. Muzi M, Spence AM, O'Sullivan F, et al. Kinetic analysis of 3′-deoxy-3′-^{18}F-fluorothymidine in patients with gliomas. *J Nucl Med.* 2006;47:1612–1621.
17. Shinomiya A, Kawai N, Okada M, et al. Evaluation of 3′-deoxy-3′-[^{18}F]-fluorothymidine (^{18}F-FLT) kinetics correlated with thymidine kinase-1 expression and cell proliferation in newly diagnosed gliomas. *Eur J Nucl Med Mol Imaging.* 2013;40:175–185.
18. Schiepers C, Chen W, Dahlbom M, Cloughesy T, Hoh CK, Huang SC. ^{18}F-fluorothymidine kinetics of malignant brain tumors. *Eur J Nucl Med Mol Imaging.* 2007;34:1003–1011.
19. Schiepers C, Dahlbom M, Chen W, et al. Kinetics of 3′-deoxy-3′-^{18}F-fluorothymidine during treatment monitoring of recurrent high-grade glioma. *J Nucl Med.* 2010;51:720–727.
20. Chen W, Delaloye S, Silverman DH, et al. Predicting treatment response of malignant gliomas to bevacizumab and irinotecan by imaging proliferation with [^{18}F] fluorothymidine positron emission tomography: a pilot study. *J Clin Oncol.* 2007;25:4714–4721.
21. Schwarzenberg J, Czernin J, Cloughesy TF, et al. 3′-deoxy-3′-^{18}F-fluorothymidine PET and MRI for early survival predictions in patients with recurrent malignant glioma treated with bevacizumab. *J Nucl Med.* 2012;53:29–36.
22. Wen PY, MacDonald DR, Reardon DA, et al. Updated response assessment criteria for high-grade gliomas: response assessment in Neuro-Oncology Working Group. *J Clin Oncol.* 2010;28:1963–1972.
23. Sullivan DC, Schwartz LH, Zhao B. The imaging viewpoint: how imaging affects determination of progression-free survival. *Clin Cancer Res.* 2013;19:2621–2628.
24. Werner-Wasik M, Nelson AD, Choi W, et al. What is the best way to contour lung tumors on PET scans? Multi-observer validation of a gradient-based method using a NSCLC digital PET phantom. *Int J Radiat Oncol Biol Phys.* 2012;82:1164–1171.
25. 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.
26. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res.* 1999;8:135–160.
27. Krak NC, Boellaard R, Hoekstra OS, Twisk JW, Hoekstra CJ, Lammertsma AA. Effects of ROI definition and reconstruction method on quantitative outcome and applicability in a response monitoring trial. *Eur J Nucl Med Mol Imaging.* 2005;32:294–301.
28. Lodge MA, Chaudhry MA, Wahl RL. Noise considerations for PET quantification using maximum and peak standardized uptake values. *J Nucl Med.* 2012;53:1041–1047.
29. Velasquez LM, Boellaard R, Kollia G, et al. Repeatability of ^{18}F-FDG PET in a multicenter phase 1 study of patients with advanced gastrointestinal malignancies. *J Nucl Med.* 2009;50:1646–1654.
30. Weber WA, Gatsonis CA, Mozley PD, et al. Repeatability of ^{18}F-FDG PET/CT in advanced non-small cell lung cancer: prospective assessment in 2 multicenter trials. *J Nucl Med.* 2015;56:1137–1143.
31. Lodge MA, Lucas JD, Marsden PK, Cronin BF, O'Doherty MJ, Smith MA. A PET study of 18FDG uptake in soft tissue masses. *Eur J Nucl Med.* 1999;26:22–30.
32. de Langen AJ, Klabbers B, Lubberink M, et al. Reproducibility of quantitative ^{18}F-3′-deoxy-3′-fluorothymidine measurements using positron emission tomography. *Eur J Nucl Med Mol Imaging.* 2009;36:389–395.
33. 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.
34. Shields AF, Lawhorn-Crews JM, Briston DA, et al. Analysis and reproducibility of 3′-deoxy-3′-[^{18}F]fluorothymidine positron emission tomography imaging in patients with non-small cell lung cancer. *Clin Cancer Res.* 2008;14:4463–4468.
35. Cole TJ. Sympercents: symmetric percentage differences on the 100 log e scale simplify the presentation of log transformed data. *Stat Med.* 2000;19:3109–3125.