

PRELIMINARY RESEARCH

Open Access

The effect of volume of interest definition on quantification of lymph node immune response to a monkeypox virus infection assessed by ^{18}F -FDG-PET

Svetlana Chefer^{1*}, Richard C Reba², Christopher Z Leyson¹, Jurgen Seidel¹, Reed F Johnson³, Joseph E Blaney³, Peter B Jahrling³ and Julie Dyal¹

Abstract

Background: 2-deoxy-2-[^{18}F]fluoro-D-glucose-positron emission tomography (^{18}F -FDG-PET) is applied in the clinic for infection assessment and is under consideration for investigating the inflammatory/immune response in lymphoid tissue in animal models of viral infection. Assessing changes in ^{18}F -FDG uptake of lymph nodes (LNs), primary lymphoid tissues targeted during viral infection, requires suitable methods for image analysis. Similar to tumor evaluation, reliable quantitation of the LN function via multiple ^{18}F -FDG-PET sessions will depend how the volume of interest is defined. Volume of interest definition has a direct effect on statistical outcome. The current study objective is to compare for the first time agreement between conventional and modified VOI metrics to determine which method(s) provide(s) reproducible standardized uptake values (SUVs) for ^{18}F -FDG uptake in the LN of rhesus macaques.

Methods: Multiple ^{18}F -FDG-PET images of LNs in macaques were acquired prior to and after monkeypox virus intravenous inoculation. We compared five image analysis approaches, SUV_{max} , SUV_{mean} , $\text{SUV}_{\text{threshold}}$, modified $\text{SUV}_{\text{threshold}}$, and $\text{SUV}_{\text{fixed volume}}$, to investigate the impact of these approaches on quantification of the changes in LN metabolic activity denoting the immune response during viral infection progression.

Results: The lowest data repeatability was observed with SUV_{max} . The best correspondence was between $\text{SUV}_{\text{fixed volume}}$ and conventional and modified $\text{SUV}_{\text{threshold}}$. A statistically significant difference in the LN ^{18}F -FDG uptake between surviving and moribund animals was shown using modified $\text{SUV}_{\text{threshold}}$ and $\text{SUV}_{\text{fixed volume}}$ (adjusted $p = 0.0037$ and $p = 0.0001$, respectively).

Conclusions: Quantification of the LN ^{18}F -FDG uptake is highly sensitive to the method applied for PET image analysis. $\text{SUV}_{\text{fixed volume}}$ and modified $\text{SUV}_{\text{threshold}}$ demonstrate better reproducibility for SUV estimates than SUV_{max} , SUV_{mean} , and $\text{SUV}_{\text{threshold}}$. $\text{SUV}_{\text{fixed volume}}$ and modified $\text{SUV}_{\text{threshold}}$ are capable of distinguishing between groups with different disease outcomes. Therefore, these methods are the preferred approaches for evaluating the LN function during viral infection by ^{18}F -FDG-PET. Validation of multiple approaches is necessary to choose a suitable method to monitor changes in LN metabolic activity during progression of viral infection.

Keywords: PET; Immune response quantitation; Animal model; SUV; Monkeypox; Intravenous inoculation

* Correspondence: schefer@mail.nih.gov

¹Division of Clinical Research, Integrated Research Facility, National Institute of Allergy and Infectious Diseases, National Institutes of Health, 8200 Research Plaza, Frederick, MD 21702, USA

Full list of author information is available at the end of the article

Background

Assessment of cell glucose metabolism by 2-deoxy-2- ^{18}F fluoro-D-glucose (^{18}F -FDG) and positron emission tomography (PET) is a powerful supplement to conventional studies of viral infection in animal models to characterize disease progression and evaluate the efficacy of potential treatments [1,2]. Lymphadenopathy is one of the predominant clinical signs of monkeypox virus infection in nonhuman primates (NHPs) and humans. Therefore, a reasonable method to monitor the evolving lymph node (LN) immune response is assessment of LN metabolic activity using standardized uptake value (SUV) as a simple semiquantitative measure of ^{18}F -FDG uptake [3,4]. Similar to the methodological issues associated with the analysis of ^{18}F -FDG-PET images in oncology [5], computing the SUVs and reliably evaluating the LN immune response to viral infection will depend on an exact and reproducible definition of the volume of interest (VOI). VOI definition for PET image quantitation is still an open research area, and users are applying the most reliable and reproducible techniques suitable for analysis of different types of disease (e.g., tumors, inflammatory conditions). Although the VOI definition is not the only factor that can affect the reproducibility of SUV estimates in the LN [5-7], the type and size of a VOI may greatly contribute to the variability of such measurements. Such variation has been previously demonstrated with tumor quantitation using ^{18}F -FDG-PET imaging [7,8].

A variety of methods have been proposed to define tumor VOI, but no reference standard has been accepted. Commonly used approaches for quantitative analysis of ^{18}F -FDG-PET images include the following: 1) measuring the value of the voxel with the highest activity within the tumor (SUV_{max}) [9,10], 2) averaging the SUVs from the voxels inside the whole tumor defined by freehand outline of tumor boundaries (SUV_{mean}) [11-13], 3) averaging the voxels with the SUVs greater than a certain percentage of SUV_{max} using thresholding techniques ($\text{SUV}_{\text{threshold}}$) [7,14,15], or 4) using fixed volume ($\text{SUV}_{\text{fixed volume}}$) defined as the average SUV within a fixed-size VOI centered over a region with high metabolic activity without conforming to the precise tumor outline. A similar concept of $\text{SUV}_{\text{fixed volume}}$ has been used by Boellaard et al. and was called SUV_{peak} [6].

These VOI metrics are rather general and, as such, are not optimized to detect reproducible changes in ^{18}F -FDG uptake (as measured by SUVs) during viral infection progression in a small target organ (LN). In particular, ^{18}F -FDG uptake during early viral infection is difficult to measure as normal or near normal LNs have low glycolytic activity comparable to background. The lack of data on the agreement between varied methods to provide reproducible SUV estimates of the metabolic activity of LN has prompted us to develop new methods for VOI

metrics. We evaluated new methods, $\text{SUV}_{\text{fixed volume}}$ and modified $\text{SUV}_{\text{threshold}}$ ($\text{mSUV}_{\text{threshold}}$), against conventional (SUV_{max} , SUV_{mean} , $\text{SUV}_{\text{threshold}}$) metrics. Here, we report the statistical reliability of each of these methods on quantitative assessment of ^{18}F -FDG uptake changes in axillary LNs of rhesus macaques (*Macaca mulatta*) following a monkeypox virus intravenous challenge. To test interscan reproducibility, data from three baseline computed tomography (CT) and PET scans prior to monkeypox virus inoculation were used to measure LN volumes and SUV_{max} , SUV_{mean} , $\text{SUV}_{\text{threshold}}$, $\text{SUV}_{\text{fixed volume}}$, and $\text{mSUV}_{\text{threshold}}$ in these animals.

Methods

Animals were housed in a facility accredited by the Association for Assessment and Accreditation of Laboratory Animal Care International. All experimental procedures were approved by the National Institute of Allergy and Infectious Diseases, Division of Intramural Research, Animal Care and Use Committee and were in compliance with the Animal Welfare Act regulations, Public Health Service policy, and the *Guide for the Care and Use of Laboratory Animals* recommendations.

Subjects

Six male rhesus macaques housed in biosafety level 3 containment, weighing 3 to 4 kg, were infected intravenously with 5×10^7 plaque forming units of monkeypox virus (MPXV Zaire 79 strain [V-79-I-005]) (for virus preparation and inoculation procedures, see Additional file 1). Three animals were treated intravenously with cidofovir (5 mg/ml/kg in Dulbecco's modified Eagle's medium; Gilead Sciences, Foster City, CA, USA) that has been shown to protect against monkeypox virus infection. The antiviral agent, cidofovir, was administered on day -1 prior to monkeypox virus challenge and on days +1, +3, +5, +7, +10, and +13 after challenge. NHPs received 25 mg/kg of probenecid by gavage 1 h before cidofovir injection to prevent cidofovir nephrotoxicity. Three animals comprised the untreated control group.

Data acquisition

Up to nine imaging sessions were conducted in each of the six animals following the procedures described previously [1]. Briefly, imaging data were acquired in animals anesthetized with isoflurane (2% to 2.5%) (Piramal Critical Care, Orchard Park, NY, USA) using a micro-PET scanner Focus-220 (Siemens AG, Malvern, PA, USA). This scanner has a bore size of 22 cm with an axial field-of-view of 7.6 cm and a transverse field-of-view of 19 cm [16]. Multiple static PET scans were initiated 1 h after the intravenous ^{18}F -FDG injection (9.25 MBq/kg) and continued for 10 min for each of two bed positions on different days over 1.5 months. Three

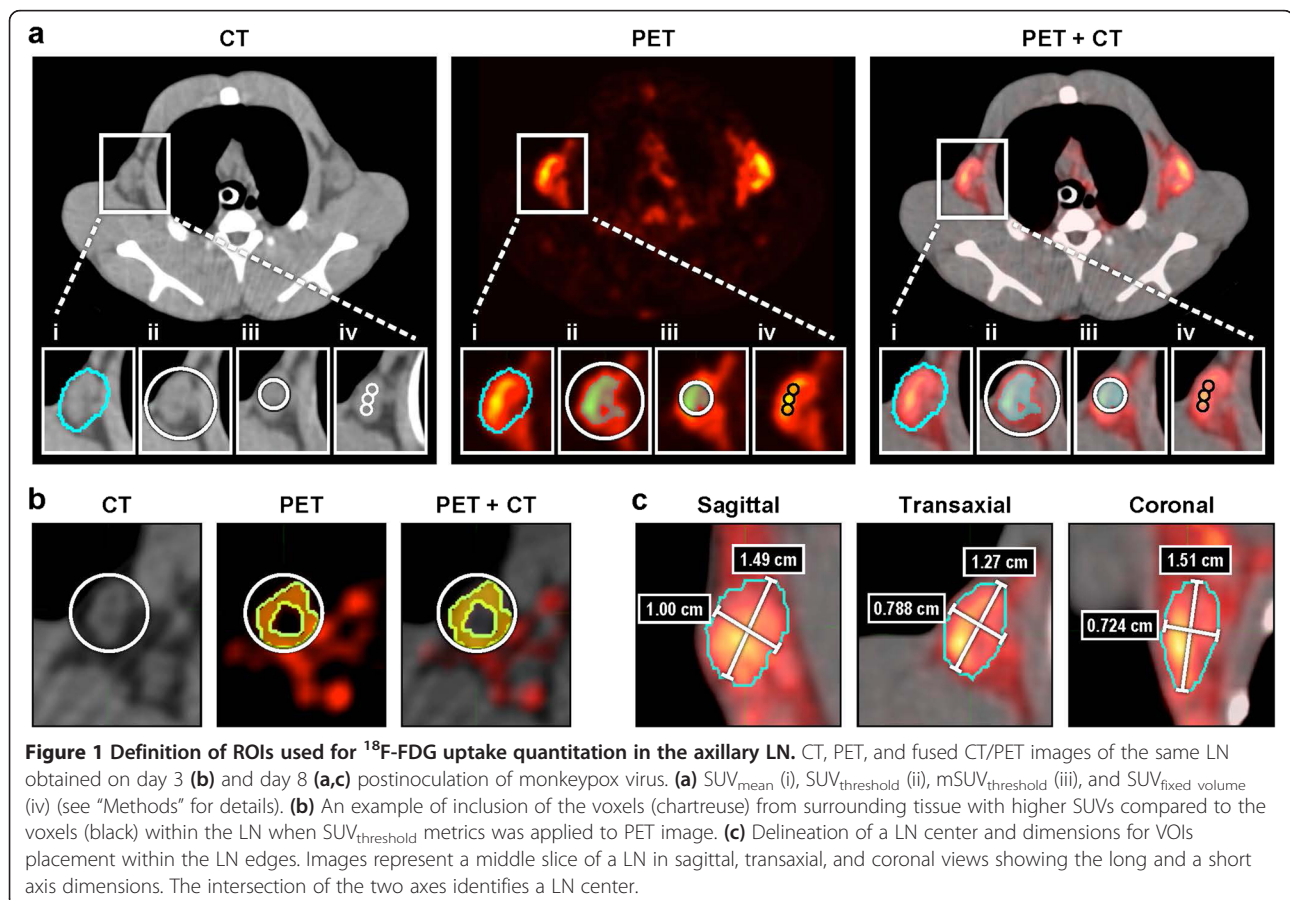
scans were performed prior to monkeypox virus inoculation (days -20, -15, and -5) and up to six scans were conducted postinoculation (days +1 or +2, +3 or +4, +7 or +8, +10, +16, and +21). The scans were conducted in the morning; the animals were fasted overnight for 12 h prior to the scanning session. The blood glucose concentrations were measured prior to the ^{18}F -FDG injection before each scanning sessions. PET images were acquired in three-dimensional (3D) mode and reconstructed iteratively using 3D-ordered subsets expectation maximization algorithm with two iterations and nine subsets followed by 18 iterations of maximum *a posteriori* reconstruction [17]. Maximum *a posteriori* parameters were adjusted to provide a uniform spatial resolution of 1.8 mm (FWHM = 1.8 mm) in all three directions. Methods for scatter, decay, random, and attenuation correction were applied during the process of PET image reconstruction.

CT images were acquired with a CereTom[®] (NeuroLogica Corp., Danvers, MA, USA) 8-slice mobile head-and-neck CT scanner that was installed in close proximity to the microPET scanner. The CereTom[®] CT scanner provided 190 slices with 0.49×0.49 mm in-plane resolution and 1.25-mm slice thickness that were acquired at 120 kVp and 5 mA. CT scans were taken either immediately

before or after PET imaging to ensure consistent animal position and fusion of the PET and CT scans for data analysis. Incorporating the use of the same table for both scanners eliminated the need for animal repositioning. To restrict animal motion, the animal was secured by anchoring the limbs and by controlling the level of anesthesia. CT scans were used for attenuation correction and coregistration with PET images to define anatomical localization of the LNs of interest. In addition, CT images were used to obtain the LN volume applied for SUV_{mean} computation to determine interscan data reproducibility prior to viral challenge.

Image analysis

Data analysis was performed using MIM workstation software version 5.2.2 (MIM Software Inc., Cleveland, OH, USA). The largest LN in the axillary fossa was chosen for analysis. This LN usually is positioned close to the body surface and easy to identify. VOIs were first specified on CT images coregistered with PET images. The volume of a whole LN was delineated by manually drawing the peripheral boundary on each of the slices where it appeared (Figure 1a (i)). SUVs from all the voxels inside this volume were averaged to calculate SUV_{mean} .



The SUV with the highest value within the LN volume identified by the MIM program was recorded as SUV_{max} .

To compute the $SUV_{threshold}$, a 'threshold' option in MIM was selected to place a sphere around the whole LN (Figure 1a (ii)). The tool averaged the voxels inside the sphere above a specified threshold that was set at 50% of the maximal voxel value. As the LN metabolic activity was often below that of surrounding tissue at baseline (Additional file 2) and during early infection (day 3/4, Additional files 2 and 3, Figure 1b), we also used a modified $SUV_{threshold}$ ($mSUV_{threshold}$) method to exclude the voxels outside the LN. $mSUV_{threshold}$ was calculated by averaging all the voxels with SUVs above 50% of the maximal value within a sphere inside the LN edges (Figure 1a (iii)). The diameter of a sphere was chosen based on the smallest LN axis defined on CT images. First, to specify the smallest axis, the CT images in 3D view were examined to identify the edges and the center of a LN (Figure 1c). When the approximate center of a LN was located, two perpendicular lines were drawn on LN slices in transaxial, coronal, and sagittal views following the largest and the smallest LN axis on each view (Figure 1c). When the sphere diameter was specified, the sphere was placed strictly within the edges of the LN (Figure 1a (iii)).

We also applied a fixed dimension method, $SUV_{fixed\ volume}$, by creating a template of three identical small spheres (0.2 cm diameter, total of 21 full voxels in all three spheres), placed contiguously within the longest axis of the LN. The sphere diameter was chosen based on the smallest LN axis (range 0.25 to 0.39 cm) among all animals determined on baseline CT images. The three spheres were transferred to each new data set by determining the center of the LN (defined by the intersection of axes in 3D view, Figure 1c), placing the middle sphere in the LN center and the other two spheres adjacent to the first one along the long axis of a LN in transaxial view. By using three small spherical VOIs, we adjusted the VOI location to the shape of the LN of each subject. Although the size and shape of the LNs differed, the VOI covered similar locations in the middle of each LN. The $SUV_{fixed\ volume}$ was computed by averaging the SUVs from 21 voxels covered by three spheres.

For SUV calculation, the radioactivity concentration from the VOI on the PET image was divided by the injected dose and normalized to the body weight of the animal and radioactive decay for the time point of ^{18}F -FDG injection

$$SUV = \frac{\text{decay corrected radioactivity concentration (MBq/ml)}}{\text{injected dose (MBq)} \div \text{body weight (g)}}$$

Statistical analysis

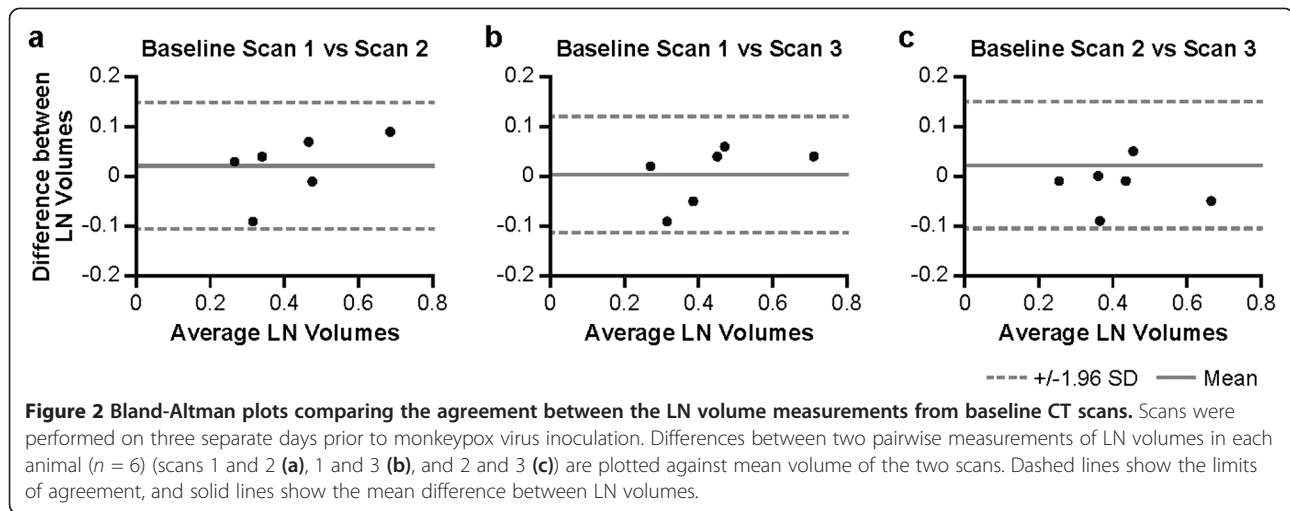
The correlation between the volume measurement on pre-inoculation CT scans 1 and 2, 1 and 3, and 2 and 3

was calculated using the Pearson product-moment correlation coefficient (r). Application of the Kolmogorov-Smirnov test [18] confirmed that the difference between the pairs of scans followed a Gaussian distribution. To investigate the interscan reproducibility of three baseline scans, we compared scans 1 and 2, 1 and 3, and 2 and 3 for LN volumes and SUV_{mean} , SUV_{max} , $SUV_{threshold}$, $mSUV_{threshold}$, and $SUV_{fixed\ volume}$ using Bland-Altman analysis [19]. The mean difference, standard deviation of the mean differences (SD), coefficient of repeatability (CR), and limits of agreement (LoA) were calculated and represented as Bland-Altman plots. The SD was calculated by squaring all the differences, adding them up, dividing them by the number of measurements, and taking the square root. The LoA were calculated by adding (upper limit) or subtracting (lower limit) the CR, defined as $CR = 1.96 \times SD$, from the mean difference. This analysis of data reproducibility was performed with the assumption that the animal health status did not change during the time the scans were obtained 5 to 21 days prior to monkeypox virus inoculation. Unchanged animal health status was confirmed by physical examination.

Bland-Altman analysis was subsequently used to investigate the agreement between five VOI metrics for peak ^{18}F -FDG uptake in the LNs of survivors on day 10 after inoculation. The differences obtained for each animal were plotted against the mean differences of the respective pairs of VOI measures. For acceptable agreement, the 95% LoA (± 1.96 SD of the mean difference) should include 95% of the difference between the methods of measurement. Two-way repeated measures analysis of variance (ANOVA) was employed to explore the difference between treated and untreated groups or surviving and moribund groups in LN ^{18}F -FDG uptake using SUVs with five different VOI metrics. We used ^{18}F -FDG uptake value at different time points (days -1, -2, -3 pre- and days +1 or +2 and +3 or +4 postinoculation with monkeypox virus) as within factor and treatment or disease outcome as between factors, respectively. *Post hoc* comparisons were performed using Bonferroni test. GraphPad Prism 6.01 (GraphPad Software Inc., La Jolla, CA, USA) was used for all statistical analyses.

Results

Three baseline CT scans prior to monkeypox virus inoculation were used to investigate interscan repeatability of analysis of LN volumes from the right axilla for application in SUV_{mean} computations. The mean LN volume and SD averaged for all six subjects from pre-inoculation scans 1, 2, and 3 were 0.43 ± 0.17 , 0.42 ± 0.14 , and 0.43 ± 0.14 cm^3 , respectively, as measured on CT images. The correlation coefficient, r , of pairwise comparisons between the LN volume determined on baseline scans 1 and 2, 1 and 3, and 2 and 3 was 0.94 ($p = 0.006$), 0.95 ($p = 0.004$),



and 0.94 ($p = 0.005$), respectively. A good agreement between the LN volume measurements on different days before inoculation was confirmed by Bland-Altman analysis (Figure 2). The mean differences and SD of the LN volume from scans 1 and 2 (Figure 2a), 1 and 3 (Figure 2b), and 2 and 3 (Figure 2c) were 0.022 ± 0.06 , 0.003 ± 0.060 , and $0.018 \pm 0.048 \text{ cm}^3$, respectively. CRs of these scans were 0.13, 0.12, and 0.09, respectively.

The range of blood glucose concentrations for all animals ($n = 6$) measured before each scanning session varied between 60 and 72 mg/dL. No correlation was noted between changes in ^{18}F -FDG uptake over the course of monkeypox virus infection and the variation in glucose concentration measured before each scanning session on different days before and after virus inoculation.

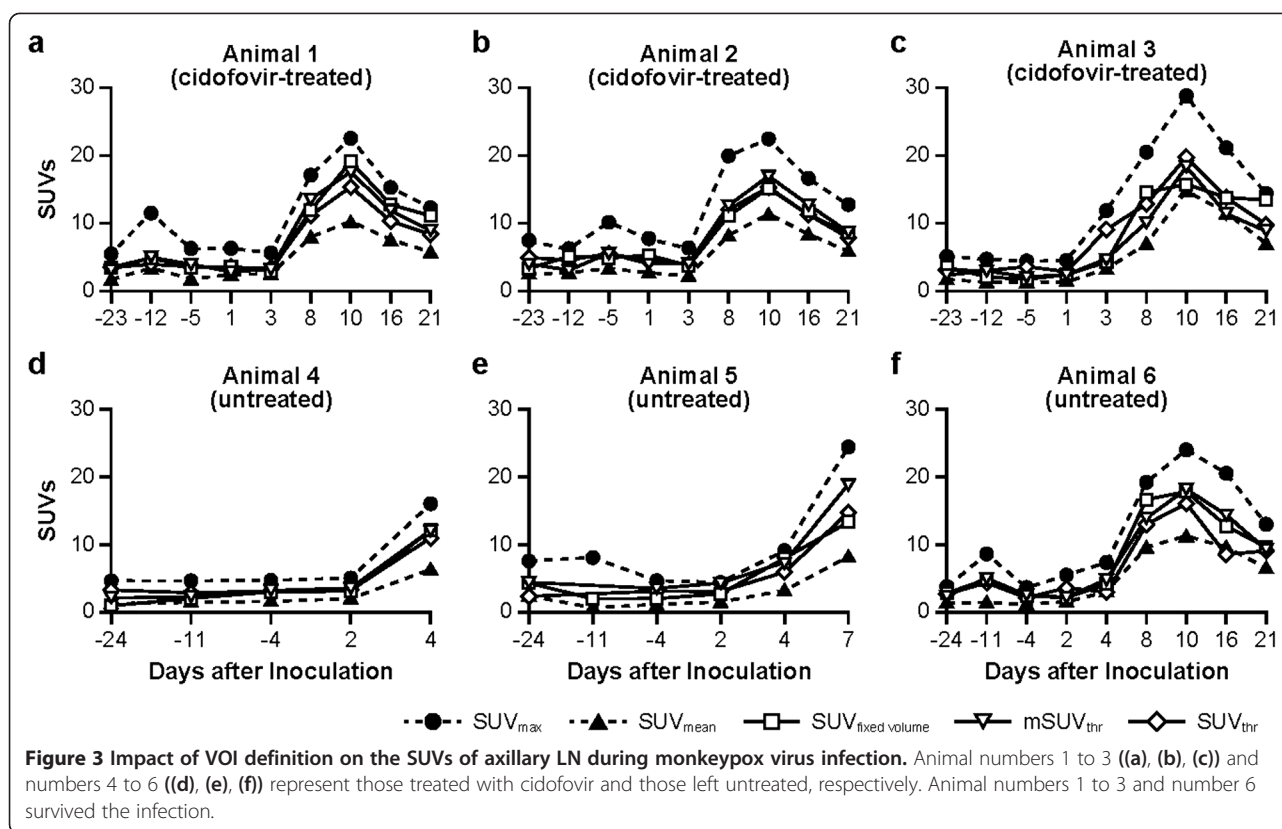
Mean SUV \pm SDs for three baseline scans, 1.98 ± 0.74 (SUV_{mean}), 6.25 ± 2.21 (SUV_{max}), 3.55 ± 0.90 (SUV_{threshold}), 3.36 ± 1.09 (mSUV_{threshold}), and 3.20 ± 1.25 (SUV_{fixed volume}) in all the animals varied depending on VOI metrics. To determine interscan variability, the mean differences in SUVs from each of the three baseline PET scans were calculated by Bland-Altman analysis. Although all the differences for the SUVs were within ± 1.96 SD, the worst agreement for the baseline scans was with SUV_{max} as shown by greater values for three outcome measures of Bland-Altman analysis, the mean difference and SD, 95% LoA, and CR (Table 1). Compared with mSUV_{threshold} and SUV_{fixed volume}, SUV_{mean} and SUV_{threshold} had smaller SDs on different scanning days as a result of larger volumes used in these SUV computations.

Following inoculation of monkeypox virus, the characterization of infection in NHPs and histological evaluation of LN tissue during infection is described in Additional file 1. One animal in the untreated group survived the infection, while the remaining two subjects became moribund on day 7 after inoculation.

In examining VOI metrics, all five SUVs showed good concurrence for the pattern of changes in LN metabolic activity in survivors (animal numbers 1 to 3 [cidofovir-treated], number 6 [untreated]) over the course of monkeypox virus infection (Figure 3). ^{18}F -FDG uptake was

Table 1 Repeatability of ^{18}F -FDG SUV measurements in the LN of normal rhesus macaques using SUVs with different VOI metrics

SUV method scan pair comparison	Mean difference \pm SD	95% limits of agreement: lower limit, upper limit	Coefficient of repeatability
SUV _{fixed volume}			
Scans 1 and 2	-0.40 ± 1.84	-4.01, 3.20	3.60
Scans 1 and 3	0.16 ± 1.74	-3.25, 3.57	3.41
Scans 2 and 3	0.57 ± 1.18	-1.75, 2.88	2.31
mSUV _{threshold}			
Scans 1 and 2	-0.89 ± 1.42	-3.68, 1.90	2.79
Scans 1 and 3	-0.41 ± 0.92	-2.21, 1.39	1.80
Scans 2 and 3	0.22 ± 1.94	-3.59, 4.03	3.81
SUV _{threshold}			
Scans 1 and 2	-0.28 ± 0.89	-2.02, 1.46	1.74
Scans 1 and 3	-0.21 ± 0.50	-1.19, 0.78	0.99
Scans 2 and 3	0.19 ± 1.26	-2.28, 2.65	2.47
SUV _{mean}			
Scans 1 and 2	0.004 ± 1.08	-2.12, 2.19	2.12
Scans 1 and 3	0.15 ± 0.72	-1.26, 1.56	1.41
Scans 2 and 3	0.12 ± 0.78	-1.41, 1.64	1.52
SUV _{max}			
Scans 1 and 2	-1.61 ± 3.01	-7.52, 4.29	5.91
Scans 1 and 3	0.004 ± 1.82	-3.52, 3.53	3.53
Scans 2 and 3	1.65 ± 3.51	-5.23, 8.54	6.88

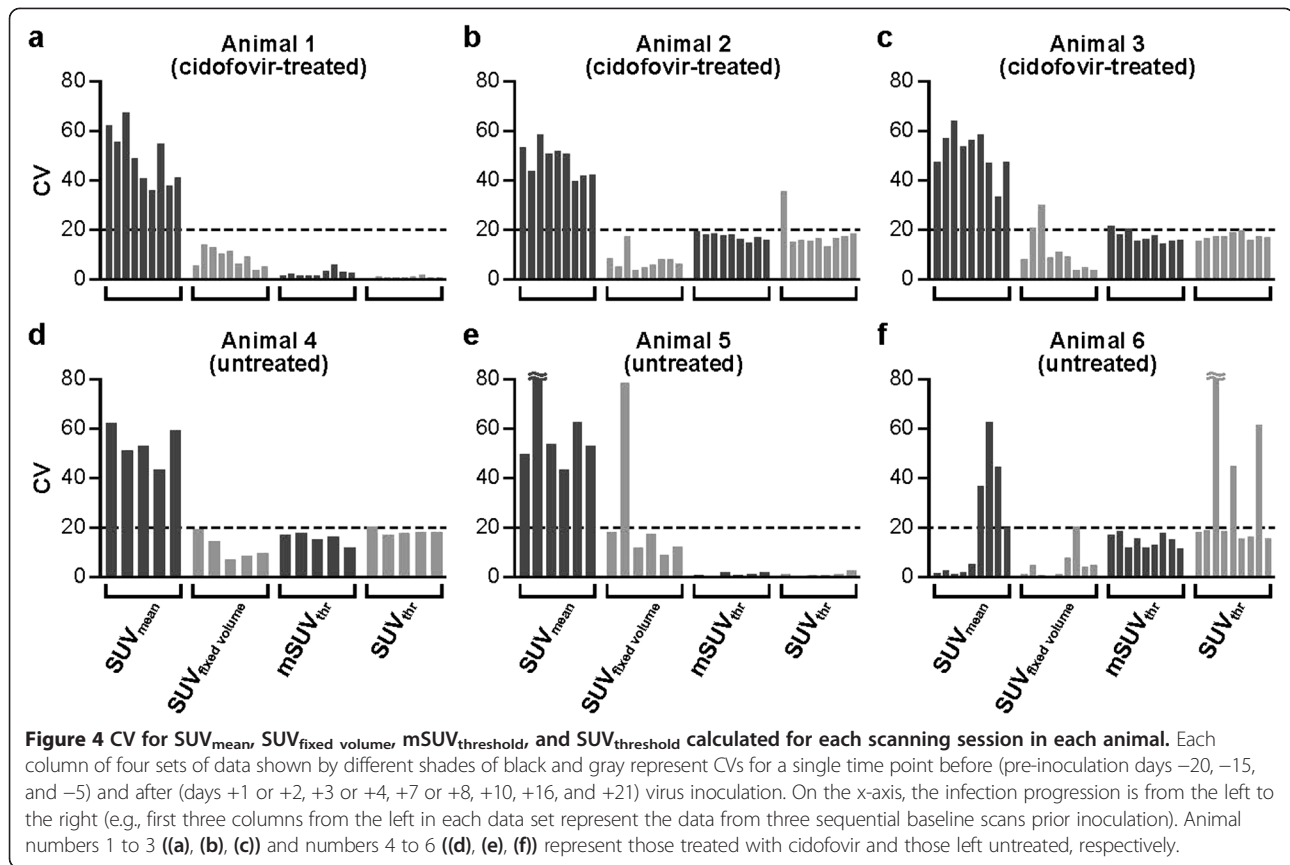


low from -5 days prior to inoculation (Additional file 3) through $+3$ days postinoculation (Additional files 2 and 3), increased markedly after day $+3$ postinoculation, and peaked on day $+10$ (Figure 3). Average SUVs \pm SDs in all surviving animals on day $+10$ postinoculation were 12.02 ± 2.00 (SUV_{mean}), 24.48 ± 3.01 (SUV_{max}), 16.65 ± 2.14 ($SUV_{threshold}$), 17.69 ± 0.69 ($mSUV_{threshold}$), and 16.98 ± 1.84 ($SUV_{fixed\ volume}$) (Figure 3a,b,c,f). Similar to the SUVs obtained from LNs prior to inoculation, the values were comparable among $SUV_{fixed\ volume}$, $mSUV_{threshold}$, and $SUV_{threshold}$ but were always above SUV_{mean} and less than the SUV_{max} by 30% to 40%. In addition, SUV_{mean} for each of the animals were characterized by substantial SD as a result of great SUV variability of the voxels included in the calculation, which led to coefficients of variation (CV) for SUV_{mean} between 40% and 60% on most scanning days (Figure 4). In contrast, the CVs for $SUV_{fixed\ volume}$, $mSUV_{threshold}$, and $SUV_{threshold}$ were rarely greater than 20%. These higher CV values were mostly observed on the images from the baseline scans with a low rate of LN metabolic activity that was comparable to surrounding tissue.

For the peak ^{18}F -FDG uptake observed in survivors on day $+10$ postinoculation, the best correspondence was between $SUV_{fixed\ volume}$ and $mSUV_{threshold}$ (Figure 5a). The mean difference and SD between the two measurements were $(-0.71) \pm 1.84$. Similar agreement was

between $mSUV_{threshold}$ and $SUV_{threshold}$ with the bias of $(1.043) \pm 1.71$ (Figure 5b) and between $SUV_{fixed\ volume}$ and $SUV_{threshold}$, but the SD was greater between the latter two SUV methods (Figure 5c). The Bland-Altman plot revealed poor agreement between $SUV_{fixed\ volume}$ and SUV_{max} , $mSUV_{threshold}$ and SUV_{max} , $SUV_{fixed\ volume}$ and SUV_{mean} , $mSUV_{threshold}$ and SUV_{mean} , and SUV_{mean} and $SUV_{threshold}$ as shown by the mean difference between any of the two pairs of 5 or greater or less than -5 (Figure 5d,e,f,g,h). Very poor agreement was found between SUV_{max} and SUV_{mean} (Figure 5i) and between SUV_{max} and $SUV_{threshold}$ (Figure 5j); the bias and SD between the two measurements were (12.5 ± 1.2) and (7.8 ± 0.89) , respectively.

PET images revealed an increase in ^{18}F -FDG-PET uptake by the LNs on day $+3$ or $+4$ postinoculation in animals that eventually became moribund compared with surviving animals (Figure 6a and Additional file 3 show CT/PET images of one representative animal from each group). To explore whether the pattern of changes in a LN metabolic activity can predict a disease outcome and specify the effect of cidofovir treatment, we analyzed a subset of the data up to day $+7$ postvirus inoculation by two-way repeated measure ANOVA. Results of statistical analysis indicate differences in the pattern of ^{18}F -FDG-PET uptake changes in moribund and surviving groups at early time points postinoculation (up to day $+3$ or $+4$)



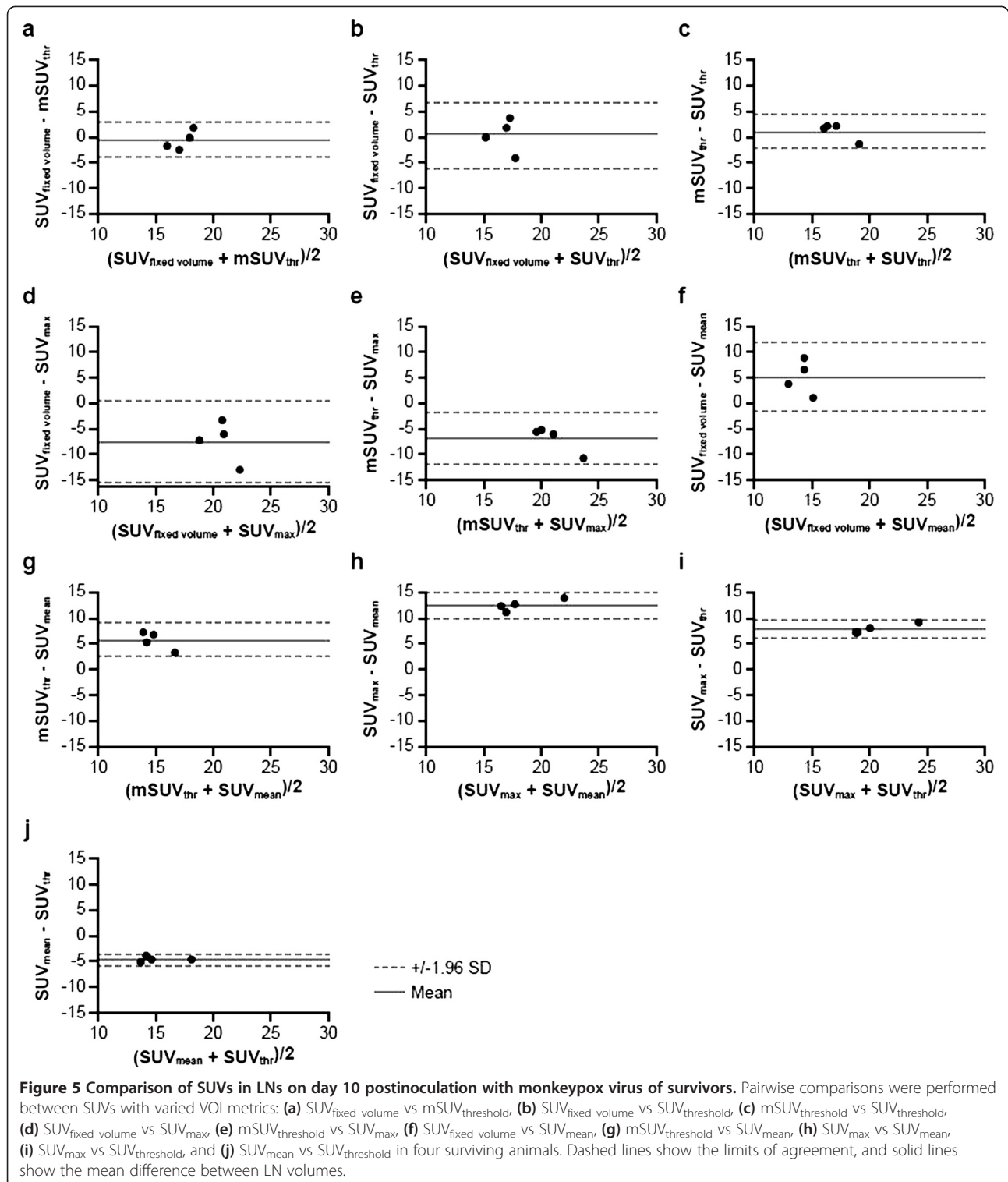
with $SUV_{fixed\ volume}$ and $mSUV_{threshold}$ only (statistically significant interaction between moribund and surviving groups, $p = 0.006$ and $p = 0.0164$, respectively, Figure 6b,c). Other SUV methods did not distinguish between changes in ^{18}F -FDG-PET uptake in moribund and surviving groups (Figure 6d,e,f). Bonferroni's multiple comparison test specified a statistically significant increase in ^{18}F -FDG uptake in moribund group on days +3 or +4 postinoculation compared with surviving group (adjusted $p = 0.0037$ and $p = 0.0001$, for $mSUV_{threshold}$ and $SUV_{fixed\ volume}$, respectively). No statistically significant effects were found between cidofovir-treated and untreated groups.

Discussion

In this paper, we compared the agreement between conventional and modified VOI metrics to identify which metric(s) provide(s) reproducible SUVs when monitoring ^{18}F -FDG uptake in the LN over the course of monkeypox virus infection (Table 2). The interscan data reproducibility for the LN volume in baseline scans ensures consistency in defining the whole LN boundaries on CT images and indicates the degree of reliability of VOI for SUV_{mean} calculation. Among five methods for the SUV calculation, SUV_{max} demonstrates the worst agreement between three baseline scans. These results confirm findings from similar studies involving cancer patients

[15,20]. Replicate scans performed within a short period of time on the same patient using an identical technique often produce poorer interstudy reproducibility for SUV_{max} than for SUVs averaging greater volumes [15]. Taking into account the low baseline rate of LN metabolic activity that was comparable to the background rate in surrounding tissue and small LN size, the repeatability of SUV measurement for the baseline scans in our study is improved by increasing the volume included in other VOI metrics.

Overall, the qualitative pattern of changes in LN ^{18}F -FDG uptake over the course of monkeypox virus infection is similar between SUVs with the five VOI metrics evaluated in this study (Figure 3). However, SUVs vary substantially depending on VOI definition. The SUVs obtained with $SUV_{threshold}$, $mSUV_{threshold}$, and $SUV_{fixed\ volume}$ VOI metrics are within similar ranges and always above the range for the SUV_{mean} and below the SUV_{max} . SUV_{max} could be the most attractive method to use for monitoring an immune response using multiple sequential PET scans because SUV_{max} is independent of the observer and simple to apply, Table 2 [10,14,23]. Despite these advantageous properties, the use of SUV_{max} is greatly influenced by adverse effects of noise [21]. This weakness of SUV_{max} in combination with a random voxel location limit the ability of SUV_{max} not only to



reproduce the measurement under normal conditions (baseline scans) but also to quantify reliably real changes in the LN metabolic activity during viral infection.

The major drawback of SUV_{mean} is the time consuming process of manual LN delineation on each of the

slices in the 3D CT images set. Similar to small tumors, LN PET images are difficult to align perfectly with CT images that often have poor soft tissue contrast. As a result, individual SUV_{mean} during infection could be underestimated by inclusion of voxels with lower metabolic

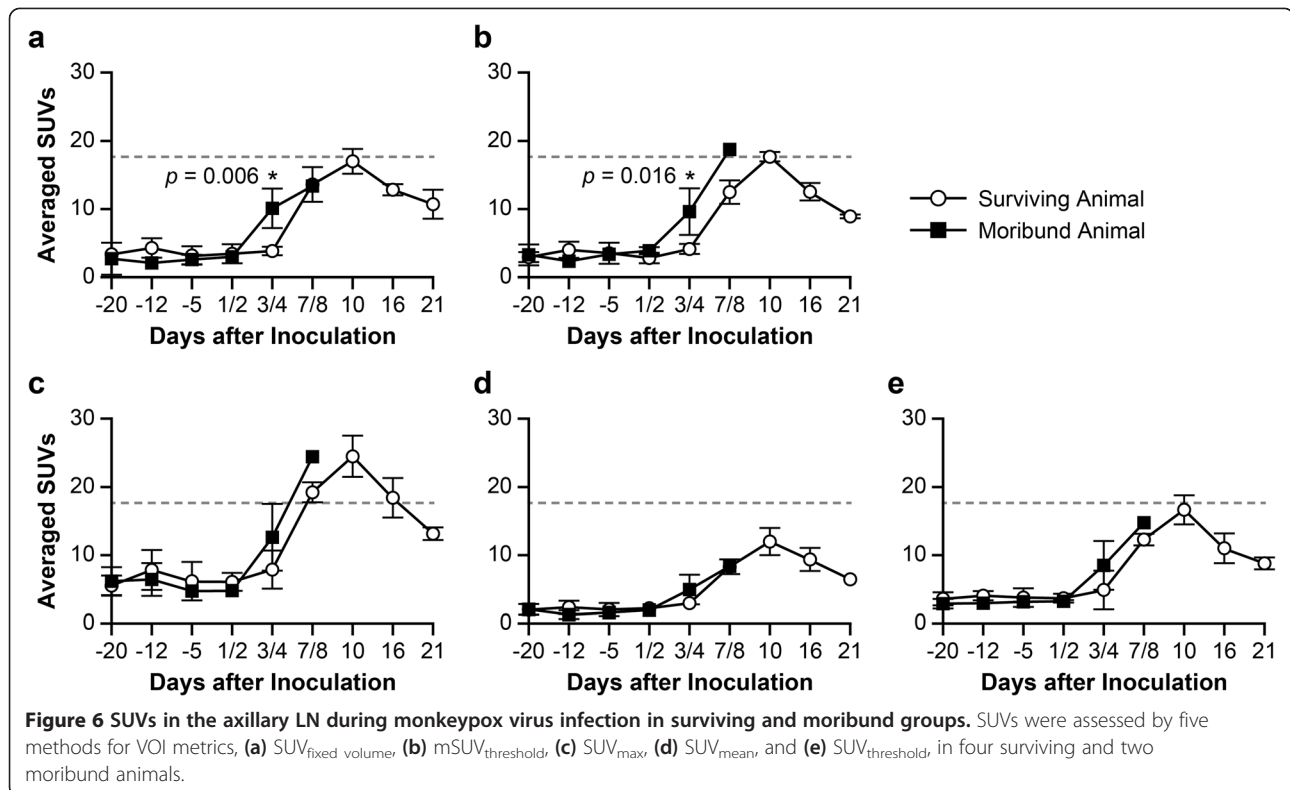


Table 2 Advantages and drawbacks of SUV methods with different VOI metrics used in the current study [5-8,11-15,21,22]

SUV metrics	Current use in clinical settings	Major drawbacks or advantages	References	Statistically significant effects in current study
SUV_{mean}	+	<ul style="list-style-type: none"> • PVE at the edges of target region • Time consuming to outline target region • Inclusion of nontarget tissue 	[5,6,11-14,21]	No
SUV_{max}	+	<ul style="list-style-type: none"> • Susceptible to noise • Random voxel location • Independent of observer • Easy to apply 	[8,21,22]	No
$SUV_{threshold}$	+	<ul style="list-style-type: none"> • Inclusion of nontarget tissue • Less sensitive to noise • Easy to apply • More reproducible than SUV_{mean} 	[7,14,15]	No
$SUV_{fixed\ volume}$ (placement adjusted in current study)	+	<ul style="list-style-type: none"> • Less sensitive to noise than SUV_{max} • Limits PVE at edges of target region • Relatively similar location of voxels between scans and between animals 	[7,8]	Yes
$mSUV_{threshold}$	-	<ul style="list-style-type: none"> • Less sensitive to noise than SUV_{max} • Relatively easy to apply • Limits PVE at edges of target region 		Yes

activity from surrounding tissue when manually defining the LN boundaries on CT images (Figure 6e). Partial-volume effect (PVE) at the edges is another factor contributing to underestimation of the values with SUV_{mean} [5,22,24]. Consequently, SUV_{mean} is characterized by substantial data variability within the VOI as shown by CVs generally greater than 40% (Figure 4). Similarly, the dynamic range of the SUV_{mean} is the widest compared with other methods (see Additional file 4).

The use of threshold-based methodology is attractive since the LN delineation on ^{18}F -FDG-PET images is easy to perform, and this method provides more reproducible measures than SUV_{max} and SUV_{mean} [7,14]. A disadvantage of the threshold technique is that the threshold is chosen rather arbitrarily, and only metabolically active tissue can be used for its application. Setting a threshold in normal LNs is more difficult compared with active LNs as metabolic activity in normal LNs appears to be similar to that in surrounding tissue (Additional files 2 and 5). At early time points postinoculation, the ^{18}F -FDG uptake in surrounding tissue is often higher than that in the LN leading to an inclusion of the majority of the voxels outside the LN edges in $SUV_{threshold}$ computation (Figure 1b, Additional file 2). Similarly, this method proved to be unsuitable to monitor tumor response after treatment since nontumor tissue is very often also included in the VOI [14]. To eliminate inclusion of voxels from surrounding tissue with high ^{18}F -FDG uptake, we limited the volume for thresholding by placing a spherical VOI in the middle of a LN. This modification in the procedure improves data variability as shown in Figure 4.

$SUV_{fixed\ volume}$ focuses on the metabolic response within a restricted location of the LN taking into account changes in LN shape and size in rhesus macaques over the course of infection. This VOI metric includes selected voxels from a slice in the middle of a LN, where metabolic activity is usually higher compared with the LN edges, and covers relatively similar locations standardized between different animals and scanning sessions. Thus, by avoiding the edges of the LN, we minimize the PVE associated with SUV_{mean} metrics, and, similar to $mSUV_{threshold}$, $SUV_{fixed\ volume}$ excludes voxels from tissues surrounding the LN. $SUV_{fixed\ volume}$ and $mSUV_{threshold}$ demonstrate the best agreement among other SUVs between the values for the peak response in survivors (Figure 6). These two methods are associated with similar statistically significant increases in ^{18}F -FDG uptake in moribund group at day +3 or +4 postinoculation of monkeypox virus compared with surviving group.

There are several limitations in our present work. Although the size of the LNs assessed in this study was within the scanner resolution, PVE correction was not applied. The best method to correct for PVE has yet to be determined as such correction by itself can produce a

bias in measured uptake. Further studies are needed to explore the relevance of PVE correction in the context of PET imaging of the LN immune response.

The aim of this preliminary study in a limited number of subjects is to compare different methods for determining the SUV and choose the method(s) with the best agreement to be applied in future characterization of a LN response to viral infection. The low number of subjects is not sufficient to demonstrate whether other more subtle effects (e.g., treatment effects), besides differentiating between disease outcomes, can be demonstrated from application of the $SUV_{fixed\ volume}$ and $mSUV_{threshold}$ metrics. The current data do not provide any explanation for the lack of difference between treated and untreated survivors in terms of ^{18}F -FDG uptake by the LN. Perhaps the lack of difference may be related to the sensitivity of the methods or the contribution of other processes not associated with metabolic activity.

Also, we cannot rule out the possibility that differences at day +3 or +4 postvirus inoculation could be explained by a suboptimal study design. To ensure consistency in viral stock properties (e.g., titers, number of passages) and in inoculation procedures, all NHPs were infected on the same day (day 0). However, imaging of treated and untreated groups was staggered over 2 days (e.g., days +3 or +4 postinoculation) to accommodate PET scanner availability and duration of pre- and postscan procedures in sick animals in a biosafety level 3 environment. Despite these limitations, a similar pattern of changes in ^{18}F -FDG uptake in the LNs is observed in treated and untreated surviving NHPs irrespective of timing of scans but is not observed between untreated surviving and moribund animals scanned on the same day. Future studies with larger group of animals and improved study design will be able to clarify this issue. In addition, an intra- and interrater reliability evaluation of LN volume and SUVs should be considered in further studies. Another point for potential criticism for the current study could be the lack of ground truth for LN ^{18}F -FDG uptake.

Conclusions

We confirmed results of previous studies that quantification of changes in ^{18}F -FDG-PET is highly sensitive to the method applied for PET image analysis. Evaluation of multiple approaches is necessary in choosing appropriate method(s) to monitor changes in LN metabolic activity during progression of infection. Results of our study indicate that $SUV_{fixed\ volume}$ and $mSUV_{threshold}$ are more reproducible than the other methods and provide the best agreement for SUV calculation. Both methods reduce the impact of noise, minimize the PVE, limit inclusion of background signals, and substantially decrease the SD of the mean SUVs. The improved precision of the SUV estimates with the proposed methods results in

statistically significant difference between moribund and surviving groups at an early stage of monkeypox virus infection that is not detected with the other three methods. Therefore, $SUV_{\text{fixed volume}}$ and $mSUV_{\text{threshold}}$ are the preferred approaches rather than SUV_{max} , $SUV_{\text{threshold}}$, or SUV_{mean} for quantitative analysis of LN immune response over the course of monkeypox virus infection using ^{18}F -FDG-PET. Consequently, these preferred methods may provide better tools for establishing ^{18}F -FDG uptake by the LNs as a marker of functional response at an early stage of monkeypox virus infection.

Additional files

Additional file 1: The macaque preparation, clinical course, and histological analyses following intravenous monkeypox virus challenge.

Additional file 2: Elevated ^{18}F -FDG uptake in tissue surrounding the LN on day 3 after virus inoculation. Maximum intensity projection (MIP) movie and representative ^{18}F -FDG-PET images fused with CT images of an axillary LN in sagittal view acquired -5 days before and day +3 and 10 after virus inoculation. High ^{18}F -FDG uptake is observed in tissue surrounding the LN on day +3 after virus inoculation.

Additional file 3: Elevated ^{18}F -FDG uptake in axillary LN of an animal that eventually became moribund. MIP movie and representative ^{18}F -FDG-PET images fused with CT images of an axillary LN in sagittal view acquired before and on day 3 after monkey virus inoculation showing elevated ^{18}F -FDG uptake in axillary LN of an animal that eventually became moribund.

Additional file 4: Dynamic range for SUV_{mean} , $SUV_{\text{fixed volume}}$, $mSUV_{\text{threshold}}$, and $SUV_{\text{threshold}}$ assessed before and after virus inoculation. Dynamic range was calculated for single time points before (pre-inoculation days -20, -15, and -5) and after (days +1 or +2, +3 or +4, +7 or +8, +10, +16, and +21) virus inoculation in each animal. On the x-axis, the infection progression is from the left to the right. Data for voxels with negative values are not included.

Additional file 5: Enlarged fused CT/PET images of axillary LN from representative moribund and surviving animals. On day -5 pre- (top row) and day +3 or +4 postvirus (bottom row) inoculation.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

JD and RFJ were involved in the study design, implementation, data collection, and manuscript preparation. RCR participated in the design of the study, data analysis, and manuscript preparation. JS and CZL were involved in the study implementation and data analysis. JEB and PJB were involved in study design. SC was involved in development of methods for image quantitation and statistical analysis and wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We thank Jennifer Hufton from the imaging team, Russell Byrum from the NIAID Comparative Medicine Branch, and Comparative Medicine veterinarians for successful implementation of PET-CT scanning protocols in the biosafety level 3 suite. In addition, we acknowledge Shen Kui for help with statistical analysis. We thank Laura Bollinger and Jiro Wada for outstanding assistance in technical writing and figure preparation of this manuscript on behalf of the Battelle Memorial Institute. This work was supported by the Division of Intramural Research of the National Institute of Allergy and Infectious Diseases (NIAID), Integrated Research Facility (NIAID), Division of Clinical Research, and Battelle Memorial Institute's prime contract with NIAID (Contract number HHS N272200700016).

Author details

¹Division of Clinical Research, Integrated Research Facility, National Institute of Allergy and Infectious Diseases, National Institutes of Health, 8200 Research Plaza, Frederick, MD 21702, USA. ²Center for Infectious Disease Imaging, Radiology and Imaging Sciences, Clinical Center, National Institutes of Health, 10 Center Drive, Bethesda, MD 20892, USA. ³Division of Intramural Research, Emerging Viral Pathogens Section, National Institute of Allergy and Infectious Diseases, National Institutes of Health, 8200 Research Plaza, Frederick, MD 21702, USA.

Received: 22 May 2014 Accepted: 4 September 2014

Published online: 16 September 2014

References

1. Dyall J, Johnson RF, Chen DY, Huzella L, Ragland DR, Mollura DJ, Byrum R, Reba RC, Jennings G, Jahrling PB, Blaney JE, Paragas J: **Evaluation of monkeypox disease progression by molecular imaging.** *J Infect Dis* 2011, **204**:1902–1911.
2. Johnson RF, Dyall J, Ragland DR, Huzella L, Byrum R, Jett C, St Claire M, Smith AL, Paragas J, Blaney JE, Jahrling PB: **Comparative analysis of monkeypox virus infection of cynomolgus macaques by the intravenous or intrabronchial inoculation route.** *J Virol* 2011, **85**:2112–2125.
3. Huang SC: **Anatomy of SUV. Standardized uptake value.** *Nucl Med Biol* 2000, **27**:643–646.
4. Thie JA: **Understanding the standardized uptake value, its methods, and implications for usage.** *J Nucl Med* 2004, **45**:1431–1434.
5. Adams MC, Turkington TG, Wilson JM, Wong TZ: **A systematic review of the factors affecting accuracy of SUV measurements.** *AJR Am J Roentgenol* 2010, **195**:310–320.
6. Boellaard R, Krak NC, Hoekstra OS, Lammertsma AA: **Effects of noise, image resolution, and ROI definition on the accuracy of standard uptake values: a simulation study.** *J Nucl Med* 2004, **45**:1519–1527.
7. 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.
8. Vanderhoek M, Perlman SB, Jeraj R: **Impact of the definition of peak standardized uptake value on quantification of treatment response.** *J Nucl Med* 2012, **53**:4–11.
9. Borst GR, Belderbos JS, Boellaard R, Comans EF, De Jaeger K, Lammertsma AA, Lebesque JV: **Standardised FDG uptake: a prognostic factor for inoperable non-small cell lung cancer.** *Eur J Cancer* 2005, **41**:1533–1541.
10. Benz MR, Evilevitch V, Allen-Auerbach MS, Eilber FC, Phelps ME, Czernin J, Weber WA: **Treatment monitoring by ^{18}F -FDG PET/CT in patients with sarcomas: interobserver variability of quantitative parameters in treatment-induced changes in histopathologically responding and nonresponding tumors.** *J Nucl Med* 2008, **49**:1038–1046.
11. Dimitrakopoulou-Strauss A, Hoffmann M, Bergner R, Uppenkamp M, Eisenhut M, Pan L, Haberkorn U, Strauss LG: **Prediction of short-term survival in patients with advanced nonsmall cell lung cancer following chemotherapy based on 2-deoxy-2-[^{18}F]fluoro-D-glucose-positron emission tomography: a feasibility study.** *Mol Imaging Biol* 2007, **9**:308–317.
12. Lucignani G: **SUV and segmentation: pressing challenges in tumour assessment and treatment.** *Eur J Nucl Med Mol Imaging* 2009, **36**:715–720.
13. Higgins KA, Hoang JK, Roach MC, Chino J, Yoo DS, Turkington TG, Brizel DM: **Analysis of pretreatment FDG-PET SUV parameters in head-and-neck cancer: tumor SUVmean has superior prognostic value.** *Int J Radiat Oncol Biol Phys* 2012, **82**:548–553.
14. Vriens D, De Geus-Oei LF, Van Laarhoven HW, Van Der Heijden HF, Krabbe PF, Visser EP, Oyen WJ: **Comparison of two region of interest definition methods for metabolic response evaluation with [(1) (8)F]FDG-PET.** *Q J Nucl Med Mol Imaging* 2010, **54**:677–688.
15. Burger IA, Huser DM, Burger C, von Schulthess GK, Buck A: **Repeatability of FDG quantification in tumor imaging: averaged SUVs are superior to SUVmax.** *Nucl Med Biol* 2012, **39**:666–670.
16. Tai YC, Ruangma A, Rowland D, Siegel S, Newport DF, Chow PL, Laforest R: **Performance evaluation of the microPET focus: a third-generation microPET scanner dedicated to animal imaging.** *J Nucl Med* 2005, **46**:455–463.
17. Qi J, Leahy RM, Cherry SR, Chatzioannou A, Farquhar TH: **High-resolution 3D Bayesian image reconstruction using the microPET small-animal scanner.** *Phys Med Biol* 1998, **43**:1001–1013.

18. Press WH: *Numerical Recipes in C: The Art of Scientific Computing*. 2nd edition. Cambridge, UK: Cambridge University Press; 1992.
19. Bland JM, Altman DG: **Statistical methods for assessing agreement between two methods of clinical measurement**. *Lancet* 1986, **1**:307–310.
20. Nahmias C, Wahl LM: **Reproducibility of standardized uptake value measurements determined by ^{18}F -FDG PET in malignant tumors**. *J Nucl Med* 2008, **49**:1804–1808.
21. Lodge MA, Chaudhry MA, Wahl RL: **Noise considerations for PET quantification using maximum and peak standardized uptake value**. *J Nucl Med* 2012, **53**:1041–1047.
22. Soret M, Bacharach SL, Buvat I: **Partial-volume effect in PET tumor imaging**. *J Nucl Med* 2007, **48**:932–945.
23. Huang YE, Chen CF, Huang YJ, Konda SD, Appelbaum DE, Pu Y: **Interobserver variability among measurements of the maximum and mean standardized uptake values on (18)F-FDG PET/CT and measurements of tumor size on diagnostic CT in patients with pulmonary tumors**. *Acta Radiol* 2010, **51**:782–788.
24. Graham M: **Quantification of Radiotracer Uptake into Tissue**. In *Molecular Imaging: Principles and Practice*. Edited by Weissleder R, Ross BD, Rehemtulla A, Gambhir SS, Shelton CT. USA: People's Medical Pub. House; 2010:1258–1270.

doi:10.1186/s13550-014-0049-z

Cite this article as: Chefer et al.: The effect of volume of interest definition on quantification of lymph node immune response to a monkeypox virus infection assessed by ^{18}F -FDG-PET. *EJNMMI Research* 2014 **4**:49.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
