OPEN ACCESS Freely available online

PLOS ONE

A Semi-Automated Technique Determining the Liver Standardized Uptake Value Reference for Tumor Delineation in FDG PET-CT

Kenji Hirata[1]*, Kentaro Kobayashi[2], Koon-Pong Wong[1], Osamu Manabe[2], Andrew Surmak[1], Nagara Tamaki[2], Sung-Cheng Huang[1]

1 Department of Molecular and Medical Pharmacology, David Geffen School of Medicine at UCLA, University of California Los Angeles, Los Angeles, California, United States of America, **2** Department of Nuclear Medicine, Hokkaido University, Graduate School of Medicine, Sapporo, Japan

Abstract

Background: ^{18}F-fluorodeoxyglucose (FDG) positron emission tomography (PET)-computed tomography (CT) has been an essential modality in oncology. We propose a semi-automated algorithm to objectively determine liver standardized uptake value (SUV), which is used as a threshold for tumor delineation.

Methods: A large spherical volume of interest (VOI) was placed manually to roughly enclose the right lobe (RL) of the liver. For each voxel in this VOI, a coefficient of variation of voxel values (CVv) was calculated for neighboring voxels within a radius of *d/2*. The voxel with the minimum CVv was then selected, where a 30-mm spherical VOI was placed at that voxel in accordance with PERCIST criteria. Two nuclear medicine physicians independently defined 30-mm VOIs manually on 124 studies in 62 patients to generate the standard values, against which the results from the new method were compared.

Results: The semi-automated method was successful in determining the liver SUV that was consistent between the two physicians in all the studies ($d = 80$ mm). The liver SUV threshold (mean +3 SD within 30-mm VOI) determined by the new semi-automated method (3.12 ± 0.61) was not statistically different from those determined by the manual method (Physician-1: 3.14 ± 0.58, Physician-2: 3.15 ± 0.58). The semi-automated method produced tumor volumes that were not statistically different from those by experts' manual operation. Furthermore, the volume change in the two sequential studies had no statistical difference between semi-automated and manual methods.

Conclusions: Our semi-automated method could define the liver SUV robustly as the threshold value used for tumor volume measurements according to PERCIST. The method could avoid possible subjective bias of manual liver VOI placement and is thus expected to improve clinical performance of volume-based parameters for prediction of cancer treatment response.

Citation: Hirata K, Kobayashi K, Wong K-P, Manabe O, Surmak A, et al. (2014) A Semi-Automated Technique Determining the Liver Standardized Uptake Value Reference for Tumor Delineation in FDG PET-CT. PLoS ONE 9(8): e105682. doi:10.1371/journal.pone.0105682

Editor: Kewei Chen, Banner Alzheimer's Institute, United States of America

Received April 30, 2014; **Accepted** July 23, 2014; **Published** August 27, 2014

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. All relevant data are within the paper and its Table S1 file.

Funding: This work was supported in part by the National Institutes of Health (NIH) grants R01-AG033015, P50-CA086306, P01-NS058489, by SNMMI Wagner-Torizuka Fellowship 2013/2015 (to Kenji Hirata), and by Hokkaido University HIROKO's Fund for Academic Exchange 2012 (to Kenji Hirata). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: kenjihirata@ucla.edu

Introduction

The clinical role of ^{18}F-fluorodeoxyglucose (FDG) positron emission tomography (PET)-computed tomography (CT) in oncology has been well established [1–3]. For prediction of treatment response, maximal standardized uptake value (SUV_{max}) has been used as the de facto standard for semi-quantitative measurement to assess the intensity of FDG uptake in the tumor [4]. Recently, an increasing number of investigators have been using volume-based parameters, such as metabolic tumor volume (MTV) and total lesion glycolysis (TLG) ($TLG = SUV_{mean}$ times MTV) for lung cancer [5,6], head-and-neck cancer [7,8], gynecological cancer [9,10], and many others. MTV is defined as the tumor volume within the boundary determined by some delineation method, such as fixed threshold (e.g. SUV≥2.5) [10–13], relative threshold (e.g. SUV≥40% of SUV_{max}) [6–9], gradient-based [5], or region-growing method [14].

Wahl et al. proposed the criteria of PERCIST 1.0 in their comprehensive review paper on FDG PET-CT for prediction of treatment response, where volume-based parameters were recommended to be obtained with the use of liver SUV as a threshold to minimize the influence of inter-study variability of tumor SUV [4]. However, manual placement of liver VOI is subjective and could still give a biased threshold and thus a biased tumor volume measurement. According to Fencl et al., a 20% change in threshold led to 20% or more change in MTV [15]. In many

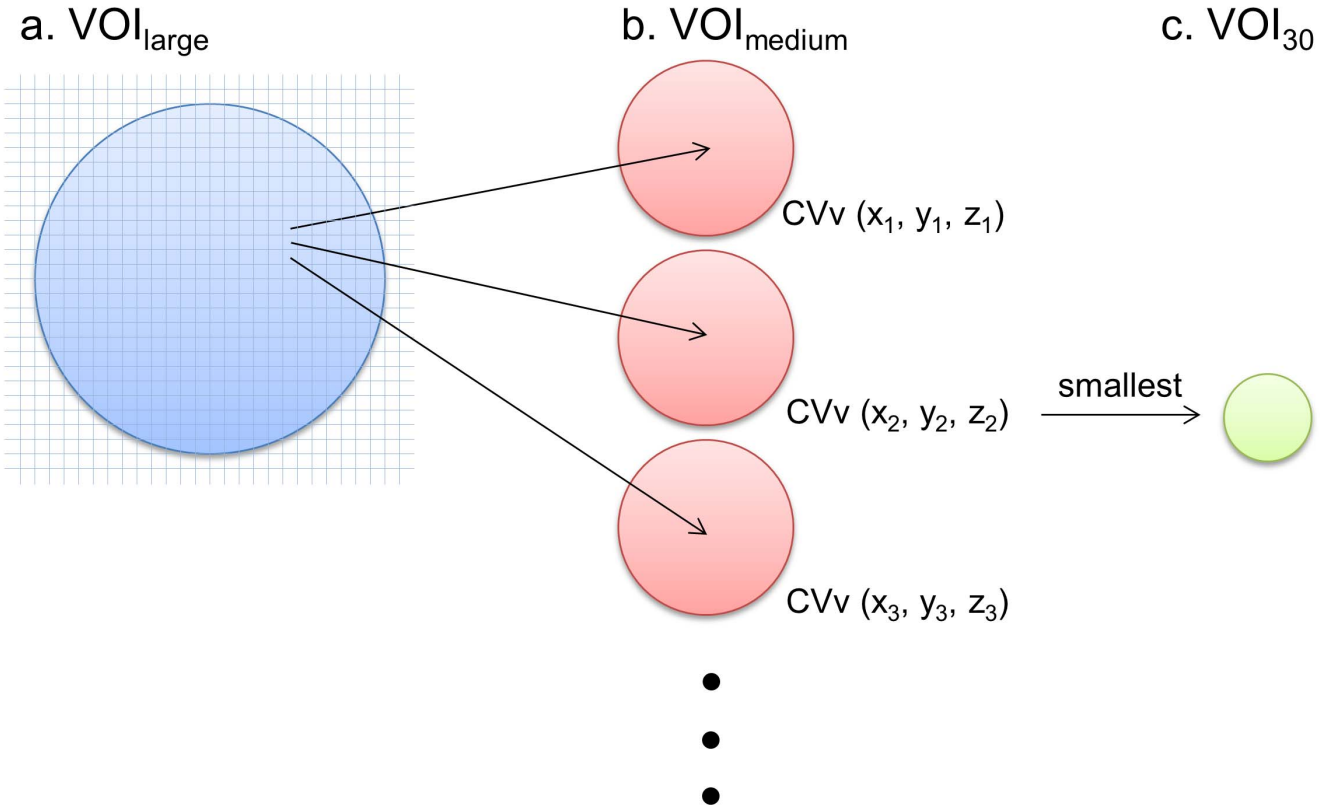

Figure 1. An illustration of the algorithm. (a) The spherical VOI_{large} confines the search area. (b) The spherical VOI_{medium} is defined for each pixel in VOI_{large} to calculate CVv (= standard deviation / mean). (c) VOI_{30} (sphere of 30-mm in diameter) is placed where CVv is smallest. The tumor delineation threshold is defined by the mean and SD within VOI_{30}.
doi:10.1371/journal.pone.0105682.g001

cases, there are heterogeneous uptakes in the liver, even without liver diseases, which could lead to variability in the resulted thresholds. Thus, an automated method for objectively placing the liver VOI for FDG PET-CT without contrast enhancement is highly desirable.

A couple of such methods have been proposed recently. Bauer et al. proposed the use of a morphological technique to identify the liver [16]. The method is simple and fast. However, the method first converts the FDG SUV image to a binary image using a fixed threshold of SUV≥1. With the inter-study variability in SUV, a fixed SUV threshold is not expected to yield reliable results. The authors realized the problem, but no specific solution has been offered. Bi et al. proposed a more complicated method with the use of both CT and PET images to automatically identify the liver [17]. However, the processing time of their proposed method is too long (>2 hours) to be practical for routine use.

In this paper, we propose a new, simple, and semi-automated method that can determine the liver VOI quickly with no inter-operator variability. In addition, the method generates threshold values comparable to those established by experts using a manual procedure. This method is expected to reduce variability in tumor volume measurements and thus improve prediction of cancer treatment response.

Methods

The new algorithm

PERCIST 1.0 criteria recommend the use of a reference value calculated from a 30-mm spherical VOI in the right lobe (RL) of the liver. The new method was designed to automate the placement of this 30-mm VOI inside RL of the liver in a reproducible fashion.

Briefly, this algorithm searches the location where the coefficient of variation associated with a voxel (CVv) of the PET images is smallest. CVv is calculated by dividing the standard deviation (SD) by the mean within a sphere of a certain diameter. Since the liver is a large organ showing moderately high FDG uptake with higher homogeneity than neighboring FDG-avid organs (e.g., colon and kidney) in the absence of pathological conditions, CVv is expected to be lower for areas inside the liver where the corresponding VOI contains only liver tissue than for marginal areas containing both liver and adjacent tissues (e.g., fat, lung, and intestines).

Fig. 1 illustrates the proposed algorithm. First, a large spherical VOI (VOI_{large}), which roughly encloses RL, is placed by an operator to confine a search area. Then, a medium-sized spherical VOI (VOI_{medium}) is defined within each pixel of the search area (i.e., the VOI_{large}) for calculation of CVv. After all CVv calculations, the voxel having the smallest CVv is identified, where a spherical VOI of 30 mm in diameter (VOI_{30}) is placed (i.e., the center of the VOI_{30} is at that voxel location). The mean and SD within VOI_{30} is used for the following image processing as the threshold of tumor delineation as suggested by PERCIST 1.0. In our implementation, the size of VOI_{large} was fixed to 150 mm in diameter. We tested different sizes of VOI_{medium}, including 40, 60, 80, 100, or 120 mm in diameter. A representative case is shown in Fig. 2.

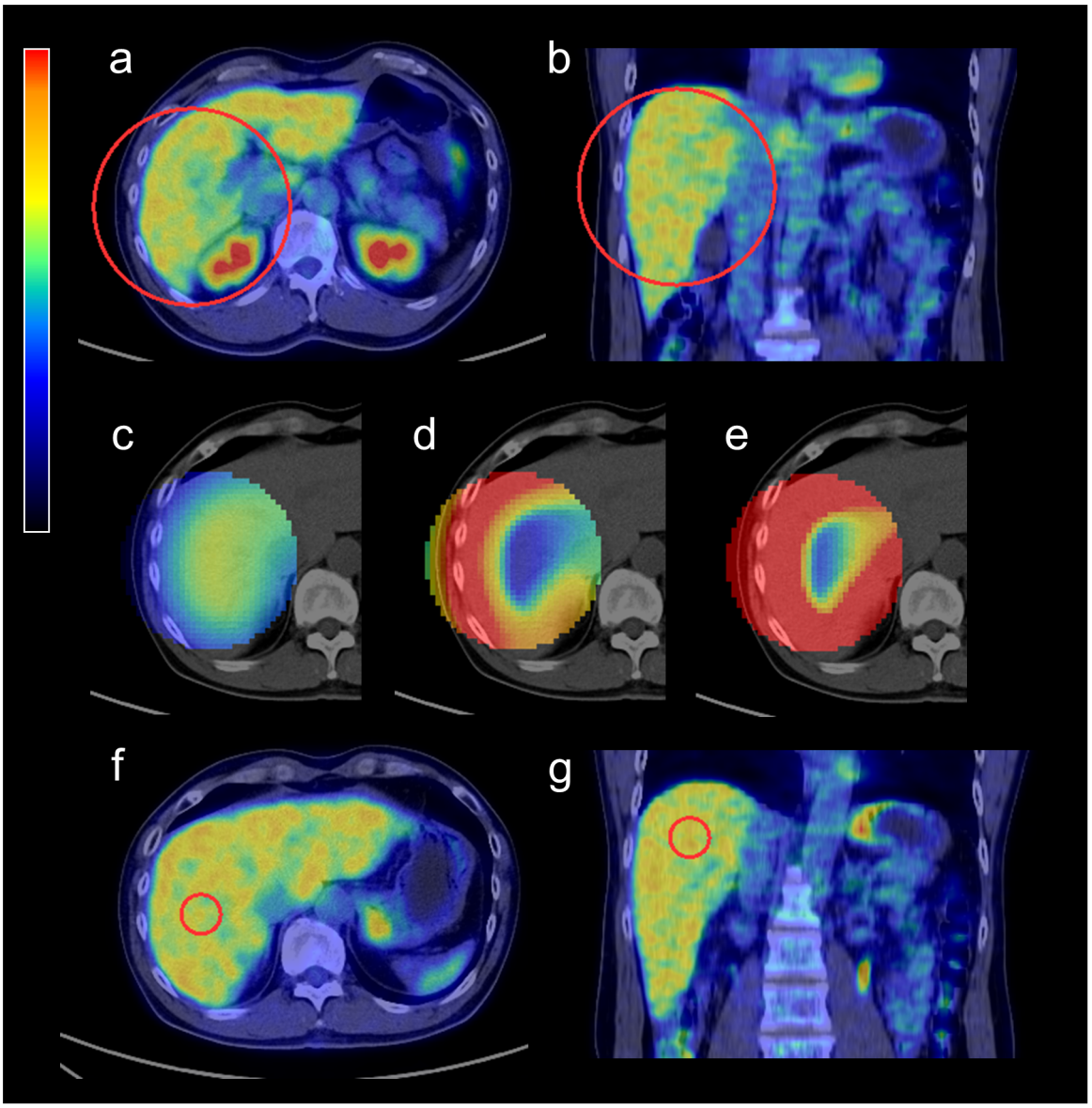

Figure 2. A representative case processed with the semi-automated algorithm. A 150-mm spherical VOI_{large} is manually placed to roughly enclose the right lobe of the liver (a, b). For each voxel within VOI_{large}, a spherical VOI_{medium} (e.g., 80 mm in diameter) is defined, and mean (c), SD (d), and coefficient of variation of voxel (CVv) (e) within VOI_{medium} are calculated. A 30-mm VOI_{30} is placed where CVv is minimized (f, g). Image color scales are 0 to 4 SUV for (a–c, f, g), 0 to 1 SUV for (d), and 0 to 0.25 (unitless number) for (e).
doi:10.1371/journal.pone.0105682.g002

The only human interaction required for this algorithm is to place VOI_{large}. The program in which we implemented the algorithm is provided, along with its usage instructions, in a public website (www.metavol.org) for free user access.

Study Subjects

Since the data were analyzed anonymously and retrospectively, no written or oral consent from each subject was obtained for the purpose of the current study. The Ethics Committee of Hokkaido University Hospital approved this retrospective study (#013-0259). A series of 1442 sequential studies of FDG PET-CT in Hokkaido University Hospital from July, 2012 to December, 2012 was reviewed. Ninety-six patients who underwent 2 or more exams were identified in this series. Thirty-one patients had 3 or more studies; in such cases, only the first two studies were included. Exclusion criteria are as follows: 1) fasting blood glucose level not measured or higher than 150 mg/dL, 2) known diabetes mellitus, 3) three or more metastatic lesions in the liver, as it becomes impossible to use liver uptake as background, 4) liver cirrhosis, 5) post-operative status of right lobectomy of the liver, and 6) significant technical error such as motion artifact and incomplete acquisition. Finally, 124 studies of 62 patients (age, 55±16 years old; body weight, 60±13 kg) were included, consisting of 24 patients with malignant lymphoma, 10 patients with head-and-neck cancer, 9 patients with skin cancers, and others. Every study was a part of patient care and performed because the patient's physician(s) considered FDG PET-CT to be necessary at that time to evaluate the patient for some clinical reasons such as staging, re-stating, and treatment response. The time interval between the two sequential examinations was 86.7±33.7 (range, 28–152) days.

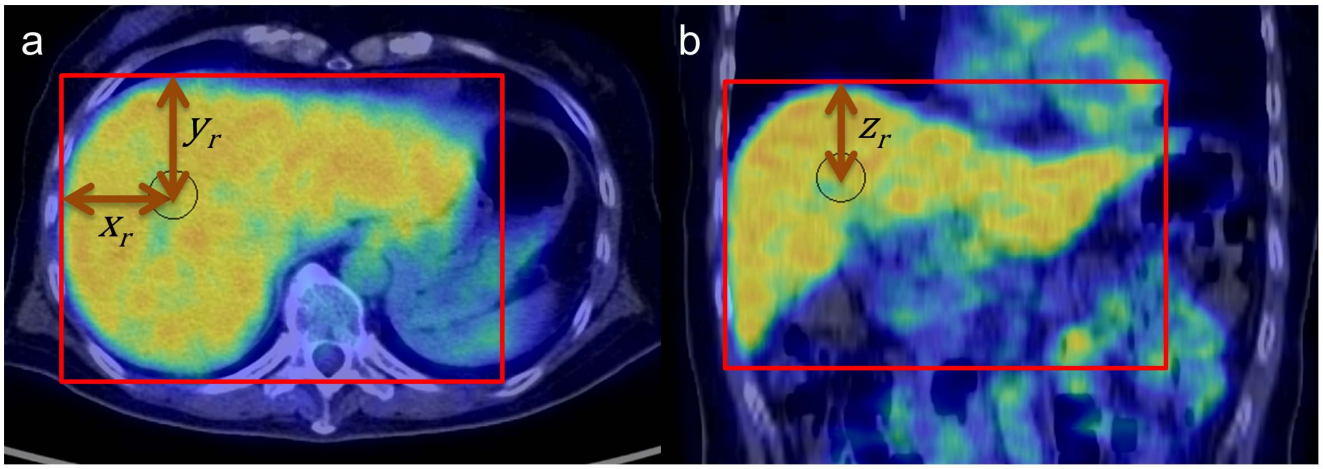

Figure 3. To evaluate inter-study same-subject reproducibility of the VOI_{30} location, a cuboid region (red rectangle) was manually created to precisely contain the whole liver. Location of VOI_{30} (black circle) was expressed as (x_r,y_r,z_r) in the *liver* coordinate system.
doi:10.1371/journal.pone.0105682.g003

Image acquisition

All clinical PET-CT studies were performed with a Biograph 64 PET-CT scanner (Asahi-Siemens Medical Technologies Ltd., Tokyo, Japan). All the patients fasted for at least 6 hours before the injection of FDG (3.7 MBq/kg), and the emission scanning was initiated 60 minutes post injection. The actual time period from FDG injection to initiation of scanning was 58.3±7.1 minutes (mean±SD). For 110 out of 124 studies (89%), the uptake time was between 50 and 70 minutes, which satisfies the PERCIST's recommendation. The transaxial and axial field of views were 58.5 cm and 21.6 cm, respectively. A 3-minute emission scanning in 3-D mode was performed for each bed position. Attenuation was corrected with X-CT images acquired without contrast media. Images were reconstructed with an iterative method integrated with point spread function (TrueX) [18]. The reconstructed image had a spatial resolution of 8.4 mm FWHM and a matrix size of 168×168 with voxel size being 4.1×4.1×2.0 mm.

Image Analysis

VOI placement. For each study, two experienced nuclear medicine physicians manually placed the 150-mm spherical VOI_{large} to roughly enclose the RL. As a control, the physicians also manually placed the 30-mm spherical VOI in RL at three different levels including upper (above the portal vein), middle (at the level of the right portal vein), and lower (below the portal vein) levels. The physicians paid careful attention not to place VOI_{30} in the area with a focally high or low uptake area.

Visual assessment of semi-automated VOI_{30}. The VOI_{30} determined by the algorithm for each study was visually categorized into either successful or failed results. The placement was considered successful when the VOI_{30} from the two physicians were located exactly in the same place inside the RL. Otherwise, the placement was considered as failure. That is, if either of the following two cases occurred, the placement was considered a "failure."

(1) VOI_{30}'s were at different locations when the two physicians ran the algorithm.

(2) VOI_{30} included non-RL region.

Table 1. Mean SUV values within-VOI_{30} from different methods and physicians.

						Significant difference
		n	Physician-1	Physician-2	ICC (95% IC)	against manual methods
	VOI level					
Manual method	Upper	124	2.51±0.47	2.53±0.47	0.975 (0.961–0.985)	
	Middle	124	2.48±0.47	2.49±0.46	0.973 (0.946–0.984)	
	Lower	124	2.47±0.48	2.48±0.51	0.970 (0.949–0.983)	
	VOI_{medium} diameter					
Semi-automated method	40-mm	119	2.52±0.49		N/A	L1, L2, M1*
	60-mm	121	2.49±0.48		N/A	None
	80-mm	124	2.49±0.48		N/A	None
	100-mm	115	2.46±0.47		N/A	U1, U2*
	120-mm	71	2.38±0.48		N/A	U1, U2*

*L1: lower level by physician-1, M1: middle level by physician-1, U1: upper level by physician-1, U2: upper level by physician-2.
doi:10.1371/journal.pone.0105682.t001

Table 2. Mean ±3 SD values within-VOI_{30} from different methods and physicians.

						Significant difference
		n	Physician-1	Physician-2	ICC (95% IC)	against manual methods
	VOI level					
Manual method	Upper	124	3.14±0.58	3.15±0.58	0.971 (0.952–0.982)	
	Middle	124	3.11±0.60	3.11±0.58	0.970 (0.953–0.979)	
	Lower	124	3.13±0.64	3.11±0.62	0.939 (0.900–0.961)	
	VOI_{medium} diameter					
Semi-automated method	40-mm	119	3.02±0.59		N/A	U1, U2, M1, M2, L1, L2*
	60-mm	121	3.09±0.59		N/A	None
	80-mm	124	3.12±0.61		N/A	None
	100-mm	115	3.15±0.61		N/A	None
	120-mm	71	3.08±0.63		N/A	M1, M2*

*L1: lower level by physician-1, L2: lower level by physician-2, *M1: middle level by physician-1, M2: middle level by physician-2, U1: upper level by physician-1, U2: upper level by physician-2.
doi:10.1371/journal.pone.0105682.t002

Inter-study same-subject reproducibility of VOI_{30} location. Since the body location of the subject varied greatly for multiple studies, we used the following coordinate system within the liver to evaluate the inter-study reproducibility of VOI_{30} placement. The nuclear medicine physician (K.H.) defined a cuboid region that precisely circumscribed the whole liver (Fig. 3). The relative location of VOI_{30} in the liver $R(x_r, y_r, z_r)$ was calculated as follows:

$$(x_r, y_r, z_r) = (x, y, z) - (x_0, y_0, z_0),$$

where (x_0, y_0, z_0) represents the VOI_{30} location in the *whole-body* coordinate, and (x_0, y_0, z_0) represents the *whole-body* coordinate values of the right, anterior, superior corner of the circumscribing cuboid. Thus, $R(x_r, y_r, z_r)$ represents the VOI_{30} location in the *liver* coordinate. Each patient underwent two PET-CT studies and had two *liver*-coordinate locations, R_1 and R_2. The Euclidean distance between R_1 and R_2 was then calculated.

Comparison of MTV. Before measuring MTV, the two physicians reached an agreement for each study about which uptake masses are considered as tumors. The tumor uptake that was proximal to non-specific uptake was not included because tumor boundary must be determined manually for these lesions, which is subjective and beyond the purpose of this study. After that, we applied mean +3 SD derived from VOI_{30} as the threshold. For all the pre-determined tumors, every voxel showing higher values than this threshold was considered as tumor. MTV was calculated as the sum of tumor volume in the entire body. Relative change was also calculated for each patient as follows:

$$\text{Relativechange} = (\text{MTV}_{\text{second}} - \text{MTV}_{\text{first}}) / \text{MTV}_{\text{first}} \times 100(\%),$$

where MTV_{first} and MTV_{second} represent the MTV derived from first and second scans, respectively. MTV and relative change were compared between two physicians.

Statistical analysis

Values are expressed as mean ±SD. The statistical software R (3.1.0) was used for all the statistical analyses. A paired t-test was used if the values could be considered as paired. The method of Holm was used to adjust the P-values for multiple comparisons [19]. Intra-class correlation (ICC) was used to evaluate inter-operator reproducibility [20]. The *psy* and *boot* packages for R were used to calculate ICC and its 95% confidential interval (CI) [21,22]. P-values less than 0.05 were considered as significant.

Results

Manually placed VOI_{30}

The mean SUV within the manually placed VOI_{30} by the first physician (P_1) vs. the second physician (P_2) from a total of 124 studies was shown in Table 1 (upper part). ICC between two physicians was highest for the upper level of the liver. Among a total of six manual values, only one combination reached

Table 3. Number of successful results by semi-automated method.

VOI_{medium} diameter (mm)	40	60	80	100	120
Successful *	119	121	124	115	71
	(96.0%)	(97.6%)	(100.0%)	(92.7%)	(57.3%)
Discrepant placement	0	0	0	0	7
Containing non-RL areas	5	3	0	9	46
Total	124	124	124	124	124

*The algorithm was considered *Successful* when the VOI_{30}'s from the two physicians were located exactly in the same place inside the RL.
doi:10.1371/journal.pone.0105682.t003

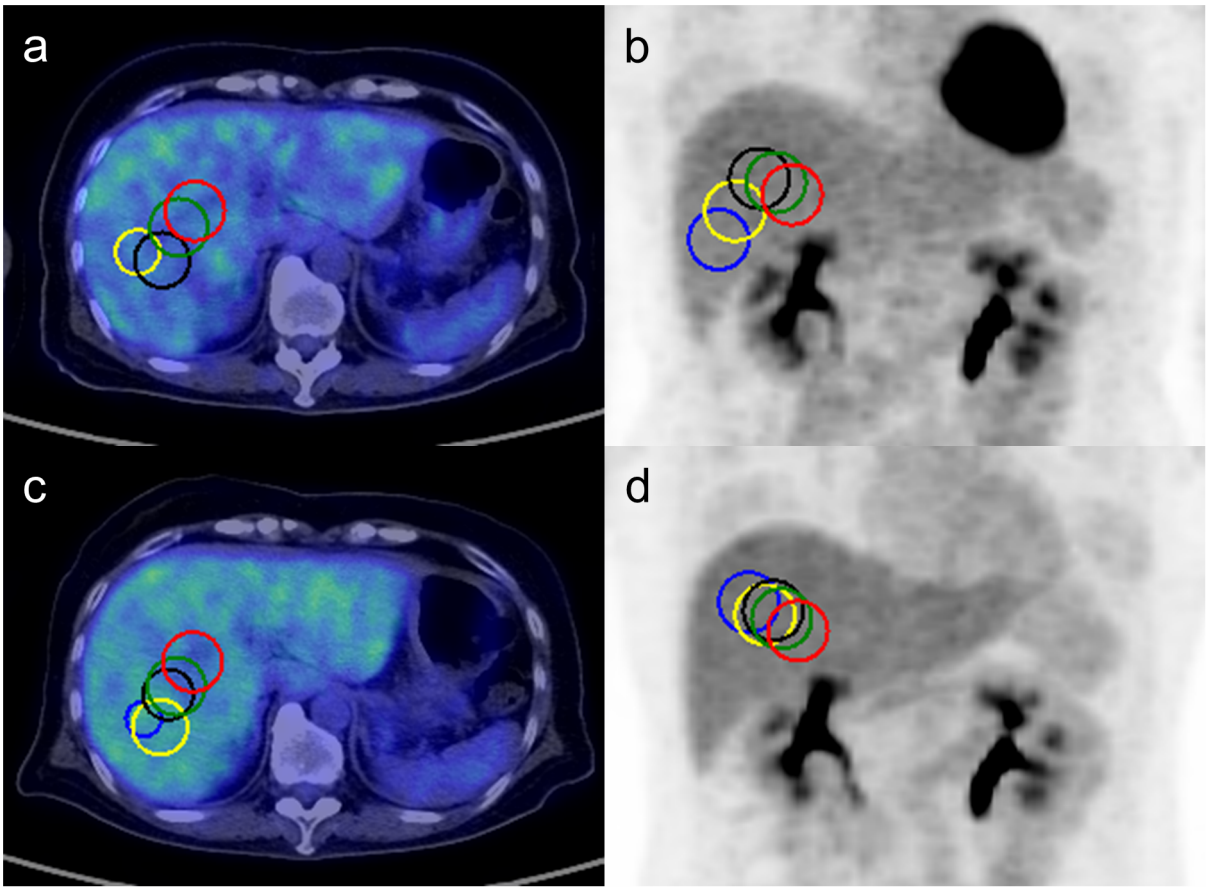

Figure 4. A representative case of two studies from the same patient. (a, b) first study and (c, d) second study. The VOI_{30}'s defined using different size of VOI_{medium} (blue: 40 mm, yellow: 60 mm, black: 80 mm, green: 100 mm, red: 120 mm) are drawn on transaxial slices (a, c) and maximum intensity projection images (b, d). The smaller VOI_{medium} located VOI_{30} further from hepatic portal region with a larger distance between the first and second studies of the same patient than larger VOI_{medium}'s did.
doi:10.1371/journal.pone.0105682.g004

significant difference (P_1's lower level vs. P_2's upper level VOI_{30}'s, $P<0.05$).

The threshold value, calculated as mean +3 SD within VOI_{30}, was shown in Table 2 (upper part). Again, ICC was highest for upper level of the liver. No significant difference was observed among six values. The results for each patient are shown in Table S1.

VOI_{30} using the semi-automated method

Visual assessment of VOI_{30} placement. Table 3 summarizes the results of visual assessment of semi-automated VOI_{30}. Using 80-mm VOI_{medium}, VOI_{30} was successfully placed in the RL for all the studies. Using a 40-mm, 60-mm, 100-mm, or 120-mm VOI_{medium}, the VOI_{30} failed to be placed appropriately in the RL in 5 (4.0%), 3 (2.4%), 9 (7.3%), and 53 (42.7%) studies, respectively. These findings suggested that the 80-mm VOI_{medium} was the best. The 120-mm VOI_{medium} was undesirable with the semi-automatic method. The computation time for 80-mm VOI_{medium} was 6.98 ± 0.02 seconds for each study (Intel Core i7-3770 CPU at 3.40 GHz).

Inter-study same-subject reproducibility of VOI_{30} location. Fig. 4 shows an example of two studies from the same patient. The automated VOI_{30} using different sizes of VOI_{medium} was superimposed on PET-CT fused images and maximum intensity projection images. The small VOI_{medium} tended to locate VOI_{30} in peripheral areas with large distances between the first and second studies of the same patient. Fig. 5 shows the Euclidian distance between VOI_{30}'s from two studies of the same patient. This analysis included only the VOI_{30} which were considered as successful at visual assessment. As the larger VOI_{medium} was used, the distance became smaller. All the combinations reached significant difference ($P<0.001$) except the combination of the 100-mm and 120-mm VOI_{medium}'s. Inter-study reproducibility was thus considered to be highest with the 100-mm and 120-mm VOI_{medium}, which was followed by the 80-mm VOI_{medium}.

Comparison of VOI_{30} values between manual vs. semi-automated methods

Tables 1 and 2 also summarize the mean of VOI_{30} (Table1, lower part) and mean +3 SD derived from VOI_{30} (Table 2, lower part), respectively. In the case that a 60-mm or 80-mm VOI_{medium} was selected, mean value of semi-automated VOI_{30} did not show significant differences from any manually derived VOI_{30}. Conversely, selection of a 40-mm, 100-mm, or 120-mm VOI_{medium} resulted in significant differences for the within-VOI_{30} mean from some manually derived VOI_{30}'s. Similarly, as Table 2 shows, in the case of 60-mm, 80-mm, or 100-mm for VOI_{medium}, threshold values from the semi-automated VOI_{30} did not show significant differences from any manually derived VOI_{30}.

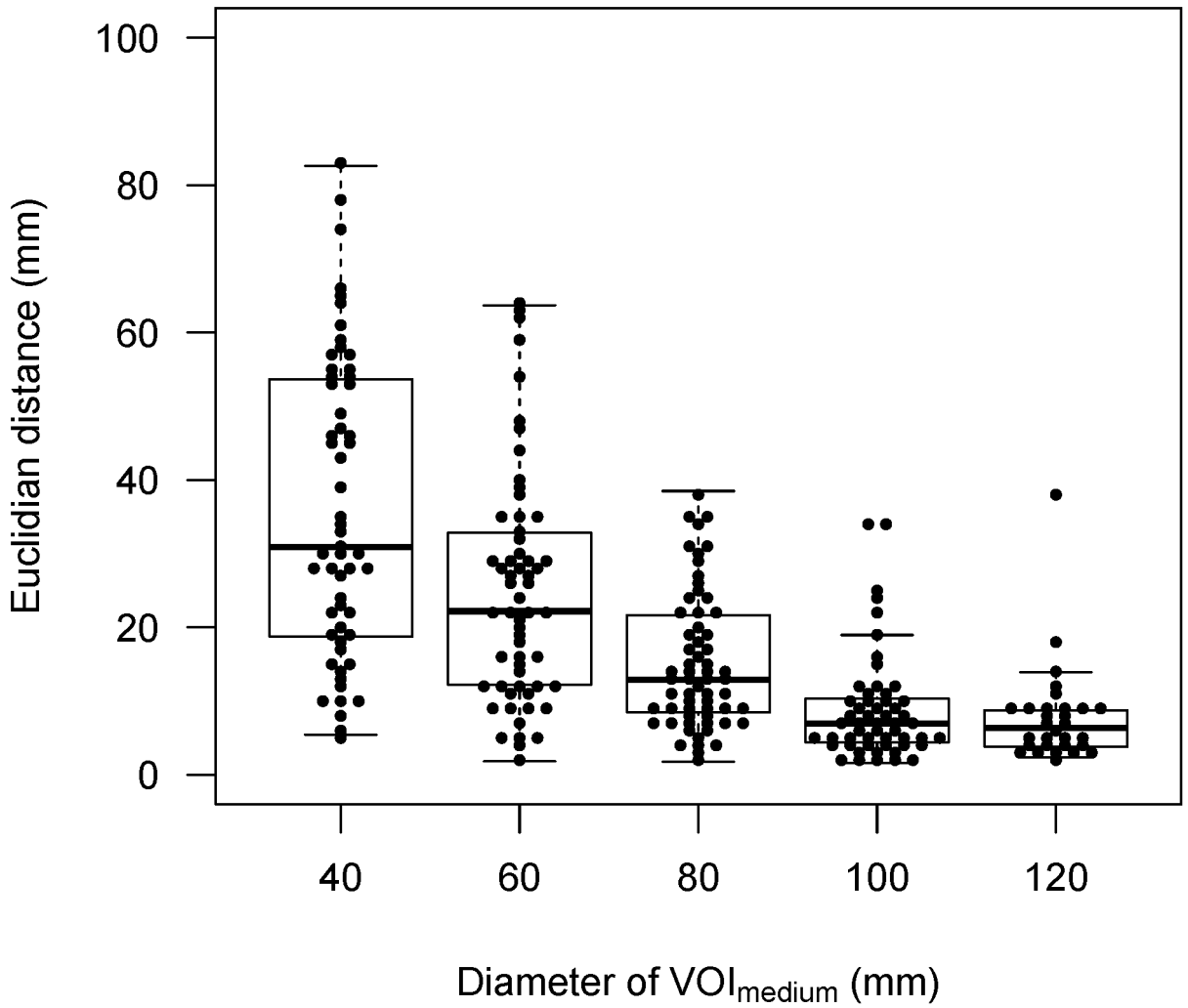

Figure 5. Euclidian distance between VOI_{30}'s from two subsequent studies of the same patient by different VOI_{medium} sizes. Except the combination of 100-mm and 120-mm, all the combinations showed significant difference ($P<0.001$) after Holm's correction for multiple comparisons.
doi:10.1371/journal.pone.0105682.g005

Overall, the 80-mm VOI_{medium} was considered to be most desirable for this new algorithm, based on the criteria of achieving (a) the best successful rate, (b) a high inter-study same-subject reproducibility of VOI_{30} location, and (c) VOI_{30} values not significantly different from those by the manual method. For the manual VOI_{30}, according to our results and previous work from another group [23], the upper VOI was considered least variable between physicians. In order to simplify the analyses, we hereafter included it only into the analysis for (i) automated VOI_{30} using 80-mm VOI_{medium}, and (ii) manual VOI_{30} placed at upper level.

MTV and its relative change

A total of 57 out of 124 studies showed measurable tumors. MTV measured using the threshold from a manually derived VOI_{30} was 49.5 ± 105.9 ml and 51.4 ± 114.5 ml from P_1 and P_2, respectively. Fig. 6a shows Bland-Altman plot between two physicians. The ICC was 0.988 (95% CI: 0.972–0.999) with bias (P_2–P_1) being 1.95 ± 16.9 ml. Paired t-test did not show a significant difference between the results by the two physicians. Using the automated algorithm, MTV was identical between the two physicians in all the patients (52.0 ± 113.8 ml, Fig. 6b). No significant bias was found between the automated and manual methods (Fig. 6c, d).

A total of 21 out of 62 patients showing measurable tumors for both first and second studies were further analyzed. The relative changes of MTV based on manually derived VOI_{30} were $-13.3\%\pm94.5\%$ and $-0.9\%\pm123.3\%$, respectively, for the two physicians (Fig. 7a). The ICC was 0.872 (95% CI: 0.603–0.997) with bias being $12.4\%\pm55.6\%$. Paired t-test did not show a significant difference between the two physicians. Using the automated algorithm, the relative change was identical between the two physicians in all the patients ($-7.9\%\pm105.3\%$, Fig. 7b). No significant bias was found between automated and manual methods (Fig. 7c, d).

Discussion

In this study, we tested a new semi-automated approach for VOI placement in RL. We applied this CV-based algorithm to clinical data and observed a high successful rate and high inter-operator reproducibility. In addition, the tumor volumes and the relative volume changes estimated using the semi-automated

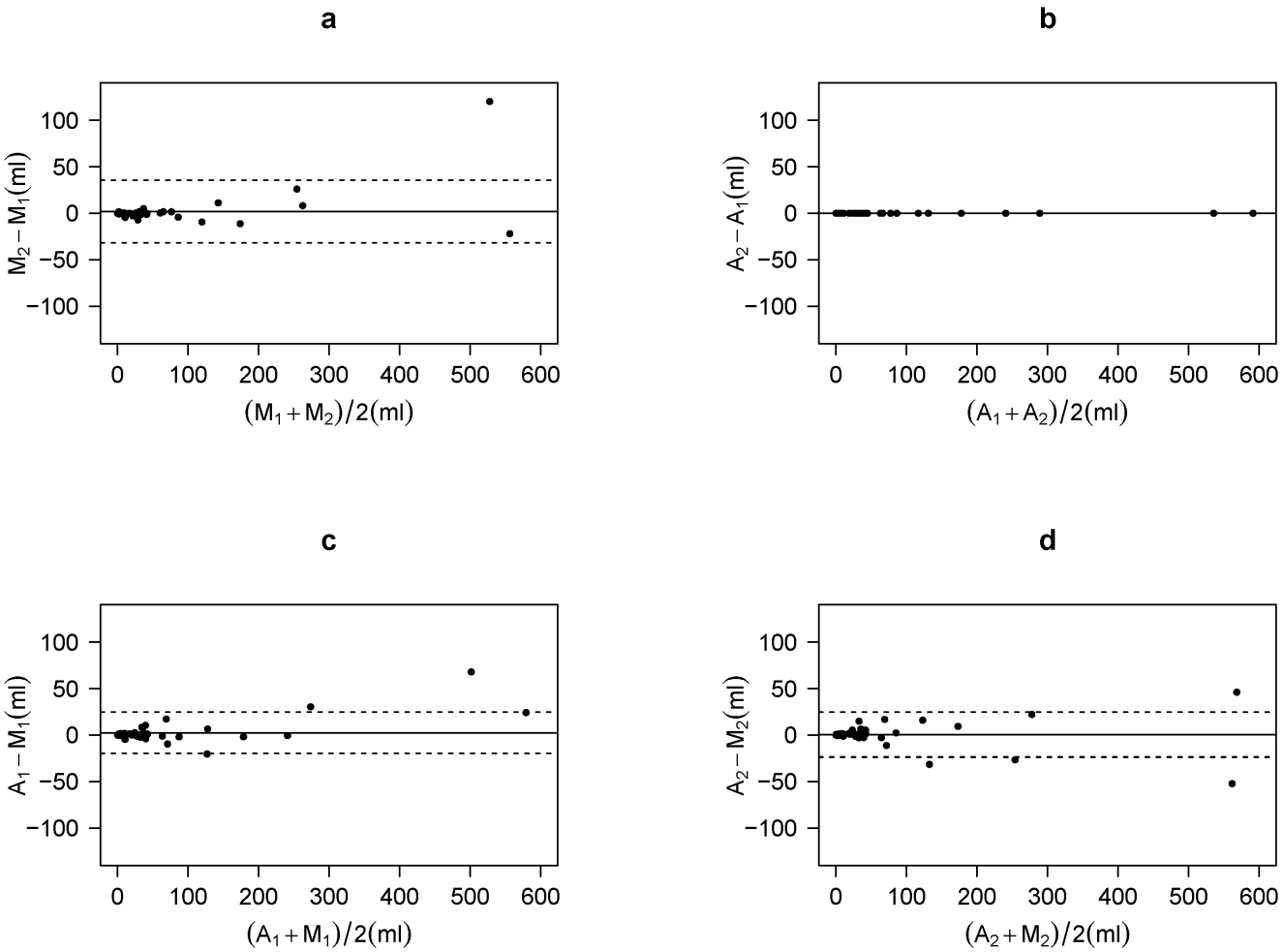

Figure 6. Bland-Altman plots of metabolic tumor volume. A_1 and A_2 represent value from the semi-automated method operated by physician-1 and -2, respectively. M_1 and M_2 represent manually derived value by physician-1 and -2, respectively. (a) M_1 vs. M_2, (b) A_1 vs. A_2, (c) A_1 vs. M_1, and (d) A_2 vs. M_2 are compared. Solid lines represent mean difference and dashed lines represent mean $\pm$2SD.
doi:10.1371/journal.pone.0105682.g006

method were not significantly different from those obtained by experts using the manual method.

Although there are existing automated methods to identify the liver for FDG PET-CT [16,17], they still need to be improved for clinical practice. On the other hand, our method is fast, simple to implement, reproducible, and highly successful (124/124 livers, 100%). Moreover, it generates equivalent values (threshold, tumor volume, and sequential volume change) to those manually determined by experts, indicating the high potential of the method in a clinical setting.

Because CVv depends on the size of the VOI_{medium}, we tried to identify the best VOI_{medium} using clinical data. First, based on successful rate with visual assessment, the 80-mm VOI_{medium} was the best with no failed results. For the other VOI_{medium}'s (i.e., 40 mm, 60 mm, 100 mm, and 120 mm in diameter), the failed cases were mainly due to localization of VOI_{30} outside RL. Next, inter-study reproducibility was tested in distance analysis, where 100-mm and 120-mm VOI_{medium}'s were the best followed by 80-mm VOI_{medium}. When a smaller VOI_{medium} was used, the CVv seemed to be a non-convex function with many local minimums found in peripheral areas of the liver, which could explain the current results (Fig. 8). Another important requirement for an automated method is, in general, that the values from the automated method should not be significantly different from the values obtained by experts. When using 80-mm VOI_{medium}, the calculated parameters (tumor volume and its relative change) had no significant difference from those determined by the manual method. All these test results support that the 80-mm VOI_{medium} should be used for this algorithm.

A recent research report showed that VOI_{30} placed in RL manually on the same FDG PET-CT study was reproducible between expert radiologists [23]. One may then argue that the automated method is not necessary. However, an expert needs to make every effort to avoid selecting focally high or low uptake areas in the liver as livers are not always homogeneous. Under the best conditions, there was still variability between operators. More importantly, one cannot rule out the possibility that an operator, intentionally or not, places the VOI_{30} in a relatively high / low uptake area in RL, yielding relatively high / low threshold value, and thus giving smaller / larger tumor volume. This might degrade the clinical value of MTV and TLG. We believe that objective methods should be used, especially for multi-center studies involving many investigators. We expect that our method will first be used in clinical trials and core. A wider experience gained in going through these uses should help determine the potential of the method for future routine clinical practice.

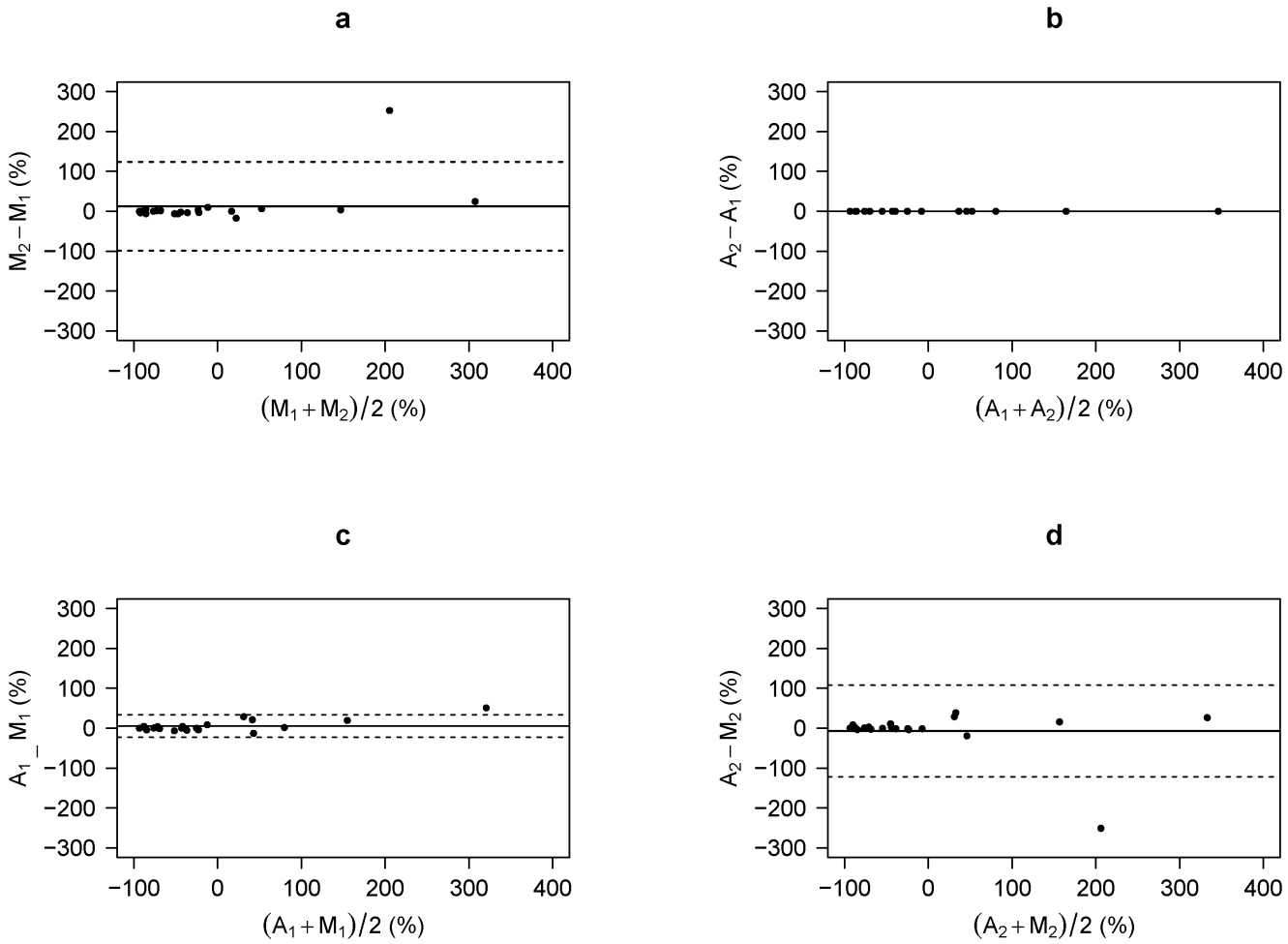

Figure 7. Bland-Altman plots of relative change of the metabolic tumor volume (MTV), calculated as [MTV_{second} – MTV_{first}] / $MTV_{first} \times 100$ (%). A_1 and A_2 represent value from the semi-automated method operated by physician-1 and -2, respectively. M_1 and M_2 represent manually derived value by physician-1 and -2, respectively. (a) M_1 vs. M_2, (b) A_1 vs. A_2, (c) A_1 vs. M_1, and (d) A_2 vs. M_2 are compared. Solid lines represent mean difference and dashed lines represent mean ±2SD.
doi:10.1371/journal.pone.0105682.g007

The 30 mm as the diameter of spherical VOI was suggested in PERCIST, but the reason for this is not very clear in the original publication. In practice, a 30-mm sphere is easy to be placed entirely within the right lobe of the liver for most patients. Although our method uses the 30-mm VOI in accordance with the suggested PERCIST guideline, the method does not require the diameter of the final VOI to be 30-mm. Therefore, the operation of the method will not be affected when future investigations determine an optimal VOI size.

PERCIST criteria have been widely used by many researchers since it was suggested in 2009 [24–29]. According to Skougaard et al., PERCIST has clear definitions and therefore more straightforward to use than conventional EORTC criteria [25]. However, even though there were studies comparing PERCIST with conventional criteria (e.g., EORTC) [25,26], PERCIST has not been proven by prospective studies to be superior in terms of diagnostic accuracy. The readers should be aware of this point when using PERCIST in clinical studies.

Use of liver VOI as a tumor boundary threshold has a shortcoming that some tumors showing relatively low FDG uptake (e.g., hepatocellular carcinoma, renal cell carcinoma, and follicular lymphoma) could be overlooked. To measure the volume of such tumors, different thresholds or more sophisticated methods will be necessary. In addition, the true threshold to distinguish active tissues from necrotic tissues is dependent on many factors, such as tumor types, tumor shapes, and partial volume effects, in addition to liver SUV.

We need to mention four limitations regarding the current study. First, we did not test the cases in which tumor FDG uptake (e.g., uterine cervical cancer) was proximal to non-tumor FDG uptake (e.g., bladder) which necessitates human interaction to delineate the tumor. An automated method to solve this issue also should be developed. Second, we did not compare the tumor volume parameters with patient outcome, although a significant difference in this aspect between the semi-automated and manual methods is not expected because the two methods did not give significant differences in the total tumor volume (Fig. 6) and in the sequential change in tumor volume (Fig. 7).

As the third limitation, the liver SUV as a reference value is appropriate only when there is no evidence of hepatic metastasis or diffuse liver diseases such as cirrhosis, or prior resection of the right lobe. Therefore, the current method works only in such cases. In liver disease cases, PERCIST recommends use of VOI placed in the blood pool of the aortic arch as an alternative, although

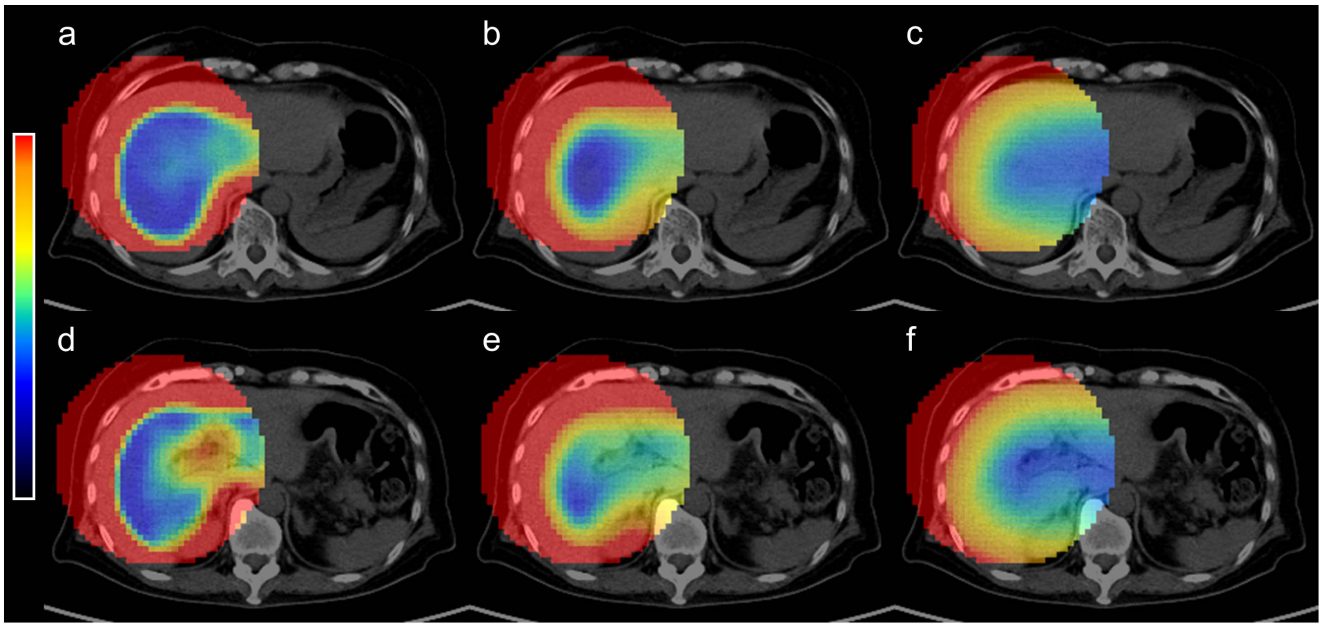

Figure 8. Example images of CVv by different VOI_{medium}'s (a, d: 40 mm; b, e: 80 mm; c, f: 120 mm). When 40-mm VOI_{medium} was used (a, d), CVv seemed to be a non-convex function with many local minimums found in peripheral areas of the liver. When 80-mm VOI_{medium} was used (b, e), CVv seemed to be a convex function. The minimum voxel existed in the right lobe of the liver. When 120-mm VOI_{medium} was used (c, f), CVv seemed to be a convex function, but the minimum existed out of the liver. Image color scales are 0 to 0.25 for (a, d), 0 to 0.50 for (b, e), and 0 to 1.00 for (c, f).
doi:10.1371/journal.pone.0105682.g008

SUV in aortic arch has a larger variability than that in the liver [4]. A previous paper suggested a fully automated method of VOI placement in the aortic arch [16]. Such a method should be helpful to minimize inter-operator variability in these cases.

Finally, we tested the semi-automated method using the images generated from only a single PET-CT scanner (i.e., Biograph 64) and with a single reconstruction algorithm (i.e., TrueX). Since our method uses CVv in the images, the results may be influenced by the degree of image noise. In addition, the study population included only Asian people. There is a possibility that other patients from different populations have different sizes of the liver, which may necessitate a different VOI_{large} and VOI_{medium}. Thus, this method needs to be evaluated further before more general uses such as in large-scale multi-center study.

Conclusions

To overcome the shortcomings of previous automated methods, we proposed a new, simple, and semi-automated method that could place consistently the 30-mm VOI in the right lobe of the liver in accordance with PERCIST 1.0 criteria. The method achieved very high successful rates and high inter-operator reproducibility, and produced tumor volumes equivalent to those through the use of manually defined VOIs by nuclear medicine experts. Avoiding subjective bias, the semi-automated method will contribute to more reliable reference values for determining tumor boundaries and, consequently, is expected to provide better prediction of cancer treatment response and prognosis.

Supporting Information

Table S1 The results of volume of interest for each patient.
(XLSX)

Acknowledgments

We thank Dr. Tomohito Kaji, Dr. Tetsuya Ishibashi, Dr. Shiro Watanabe, Dr. Takuya Toyonaga, and Dr. Keiichi Magota for their supports.

Author Contributions

Conceived and designed the experiments: KH SH. Performed the experiments: KH KK. Analyzed the data: KH. Contributed reagents/materials/analysis tools: KH KW OM. Contributed to the writing of the manuscript: KH KW OM AS NT SH.

References

1. Ben-Haim S, Ell P (2009) 18F-FDG PET and PET/CT in the evaluation of cancer treatment response. J Nucl Med 50: 88–99.
2. Jhanwar YS, Straus DJ (2006) The role of PET in lymphoma. J Nucl Med 47: 1326–1334.
3. Israel O, Kuten A (2007) Early detection of cancer recurrence: 18F-FDG PET/CT can make a difference in diagnosis and patient care. J Nucl Med 48 Suppl 1: 28S–35S.
4. Wahl RL, Jacene H, Kasamon Y, Lodge MA (2009) From RECIST to PERCIST: Evolving Considerations for PET response criteria in solid tumors. J Nucl Med 50 Suppl 1: 122S–150S.
5. Liao S, Penney BC, Wroblewski K, Zhang H, Simon CA, et al. (2012) Prognostic value of metabolic tumor burden on 18F-FDG PET in nonsurgical patients with non-small cell lung cancer. Eur J Nucl Med Mol Imaging 39: 27–38.
6. Chen HH, Chiu NT, Su WC, Guo HR, Lee BF (2012) Prognostic value of whole-body total lesion glycolysis at pretreatment FDG PET/CT in non-small cell lung cancer. Radiology 264: 559–566.
7. Ryu IS, Kim JS, Roh JL, Lee JH, Cho KJ, et al. (2013) Prognostic value of preoperative metabolic tumor volume and total lesion glycolysis measured by 18F-FDG PET/CT in salivary gland carcinomas. J Nucl Med 54: 1032–1038.

8. Chu KP, Murphy JD, La TH, Krakow TE, Iagaru A, et al. (2012) Prognostic value of metabolic tumor volume and velocity in predicting head-and-neck cancer outcomes. Int J Radiat Oncol Biol Phys 83: 1521–1527.
9. Kidd EA, Thomas M, Siegel BA, Dehdashti F, Grigsby PW (2013) Changes in cervical cancer FDG uptake during chemoradiation and association with response. Int J Radiat Oncol Biol Phys 85: 116–122.
10. Liao S, Lan X, Cao G, Yuan H, Zhang Y (2013) Prognostic predictive value of total lesion glycolysis from 18F-FDG PET/CT in post-surgical patients with epithelial ovarian cancer. Clin Nucl Med 38: 715–720.
11. Byun BH, Kong CB, Park J, Seo Y, Lim I, et al. (2013) Initial Metabolic Tumor Volume Measured by 18F-FDG PET/CT Can Predict the Outcome of Osteosarcoma of the Extremities. J Nucl Med 54: 1725–1732.
12. Usmanij EA, de Geus-Oei LF, Troost EG, Peters-Bax L, van der Heijden EH, et al. (2013) 18F-FDG PET early response evaluation of locally advanced non-small cell lung cancer treated with concomitant chemoradiotherapy. J Nucl Med 54: 1528–1534.
13. Bazan JG, Koong AC, Kapp DS, Quon A, Graves EE, et al. (2013) Metabolic tumor volume predicts disease progression and survival in patients with squamous cell carcinoma of the anal canal. J Nucl Med 54: 27–32.
14. Kao CH, Lin SC, Hsieh TC, Yen KY, Yang SN, et al. (2012) Use of pretreatment metabolic tumour volumes to predict the outcome of pharyngeal cancer treated by definitive radiotherapy. Eur J Nucl Med Mol Imaging 39: 1297–1305.
15. Fencl P, Belohlavek O, Harustiak T, Zemanova M (2012) The analysis of factors affecting the threshold on repeated 18F-FDG-PET/CT investigations measured by the PERCIST protocol in patients with esophageal carcinoma. Nucl Med Commun 33: 1188–1194.
16. Bauer C, Sun S, Sun W, Otis J, Wallace A, et al. (2012) Automated measurement of uptake in cerebellum, liver, and aortic arch in full-body FDG PET/CT scans. Med Phys 39: 3112–3123.
17. Bi L, Kim J, Wen L, Feng DD (2012) Automated and robust PERCIST-based thresholding framework for whole body PET-CT studies. Conf Proc IEEE Eng Med Biol Soc 2012: 5335–5338.
18. Panin VY, Kehren F, Michel C, Casey M (2006) Fully 3-D PET reconstruction with system matrix derived from point source measurements. IEEE Trans Med Imaging 25: 907–921.
19. Holm S (1979) A simple sequentially rejective multiple test procedure. Scand J Stat 6: 65–70.
20. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33: 159–174.
21. Falissard B (2012) psy: Various procedures used in psychometry.
22. Davison ACH, Hinkley DV (1997) Bootstrap Methods and Their Applications.: Cambridge University Press, Cambridge.
23. Viner M, Mercier G, Hao F, Malladi A, Subramaniam RM (2013) Liver SULmean at FDG PET/CT: interreader agreement and impact of placement of volume of interest. Radiology 267: 596–601.
24. Mhlanga JC, Durand D, Tsai HL, Durand CM, Leal JP, et al. (2014) Differentiation of HIV-associated lymphoma from HIV-associated reactive adenopathy using quantitative FDG PET and symmetry. Eur J Nucl Med Mol Imaging 41: 596–604.
25. Skougaard K, Nielsen D, Jensen BV, Hendel HW (2013) Comparison of EORTC criteria and PERCIST for PET/CT response evaluation of patients with metastatic colorectal cancer treated with irinotecan and cetuximab. J Nucl Med 54: 1026–1031.
26. Thacker CA, Weiss GJ, Tibes R, Blaydorn L, Downhour M, et al. (2012) 18-FDG PET/CT assessment of basal cell carcinoma with vismodegib. Cancer Med 1: 230–236.
27. Gilbert J, Rudek MA, Higgins MJ, Zhao M, Bienvenu S, et al. (2012) A phase I trial of erlotinib and concurrent chemoradiotherapy for stage III and IV (M0) squamous cell carcinoma of the head and neck. Clin Cancer Res 18: 1735–1742.
28. Tateishi U, Miyake M, Nagaoka T, Terauchi T, Kubota K, et al. (2012) Neoadjuvant chemotherapy in breast cancer: prediction of pathologic response with PET/CT and dynamic contrast-enhanced MR imaging–prospective assessment. Radiology 263: 53–63.
29. Yanagawa M, Tatsumi M, Miyata H, Morii E, Tomiyama N, et al. (2012) Evaluation of response to neoadjuvant chemotherapy for esophageal cancer: PET response criteria in solid tumors versus response evaluation criteria in solid tumors. J Nucl Med 53: 872–880.