

# Standardized added metabolic activity (SAM): a partial volume independent marker of total lesion glycolysis in liver metastases

Jeroen Mertens · André Dobbeleir · Hampshire Ham ·  
Yves D'Asseler · Ingeborg Goethals ·  
Christophe Van de Wiele

Received: 9 January 2012 / Accepted: 22 May 2012 / Published online: 15 June 2012  
© Springer-Verlag 2012

## Abstract

**Purpose** The standardized added metabolic activity (SAM) is a new marker of total lesion glycolysis that avoids partial volume effect (PVE) and thresholding. SAM is calculated by drawing a volume of interest (VOI<sub>1</sub>) around the tumour and a larger VOI (VOI<sub>2</sub>) around VOI<sub>1</sub>. Subtracting the background activity in VOI<sub>2</sub>-VOI<sub>1</sub> from VOI<sub>1</sub> yields SAM. If VOI<sub>1</sub> is set at a reasonable distance from the tumour, PVE are avoided. Phantom and initial clinical validation data are presented.

**Methods** Spheres of a Jaszczak phantom were filled with a 5.4, 3.64 and 2.0 times higher concentration relative to background activity and positron emission tomography (PET) data were acquired during 10 min. SAM of all spheres was expressed as a percentage of the expected value (the actual activity ratio minus 1). In 15 patients a 10-min list-mode acquisition PET study centred on their primary squamous cell carcinoma (PSCC) was performed and images of 1-10 min reconstructed. SAM<sub>1-9min</sub> values of PSCC were expressed as a percentage of SAM<sub>10min</sub>. Nineteen patients suffering from liver metastases treated with chemotherapy underwent PET/CT prior to (scan 1) and after 3–6 cycles of chemotherapy (scan 2). SAM and maximum standardized uptake values (SUV<sub>max</sub>) of the liver lesions on scan 1 (SAM<sub>1</sub> and SUV<sub>max1</sub>) and the percentage reduction between both  $\Delta$ SAM and

$\Delta$ SUV<sub>max</sub> were related to Response Evaluation Criteria in Solid Tumors (RECIST) response.

**Results** For the phantom acquisitions, the mean normalized SAM/sphere volume calculated was 94.9 % (SD 5.9 %) of the expected value. In the PSCC patients, the mean difference between SAM<sub>1min</sub> and SAM<sub>10min</sub> was only 4 % (SD 5 %). SUV<sub>max1min</sub> and SUV<sub>max10min</sub> proved to be not significantly different, but the variability was slightly larger than that of SAM (SD 6.4 %). SAM<sub>1</sub> and  $\Delta$ SAM values for responders versus non-responders were, respectively, 57 (SD 119) versus 297 (SD 625) for SAM<sub>1</sub> ( $p=0.2$ ) and 99 % (SD 3 %) versus 32 % (SD 44 %) for  $\Delta$ SAM ( $p=0.001$ ). SUV<sub>max1</sub> and  $\Delta$ SUV<sub>max</sub> values in responders versus non-responders were, respectively, 3.9 (SD 2.4) versus 6.3 (SD 3.1) for SUV<sub>max1</sub> ( $p=0.08$ ) and 94 % (SD 17) versus 7 % (SD 40 %) for  $\Delta$ SUV<sub>max</sub> ( $p=0.0001$ ). The AUC of  $\Delta$ SAM and  $\Delta$ SUV<sub>max</sub> were not significantly different on receiver-operating characteristic (ROC) analysis (AUC 1.0 and 0.99, respectively,  $p=0.6$ ).

**Conclusion** SAM is a promising parameter for tumour response assessment of liver metastases by means of <sup>18</sup>F-fluorodeoxyglucose PET.

**Keywords** Standardized added metabolic activity · Partial volume effect · Therapy monitoring · Total lesion glycolysis · <sup>18</sup>F-FDG PET

J. Mertens · A. Dobbeleir · H. Ham · Y. D'Asseler · I. Goethals ·  
C. Van de Wiele (✉)  
Department of Nuclear Medicine, Ghent University Hospital,  
De Pintelaan 185,  
9000 Ghent, Belgium  
e-mail: Christophe.VandeWiele@ugent.be

J. Mertens  
e-mail: Jeroen.Mertens@ugent.be

## Introduction

Standardized uptake values (SUVs) or metabolic rates are routinely used to monitor response to therapy when using <sup>18</sup>F-fluorodeoxyglucose (FDG) positron emission tomography (PET). However, both parameters incompletely capture the

tumour response to therapy as they do not reflect the total tumour metabolic activity but instead reflect metabolic activity per gram of tissue; a tumour could grow or shrink without a change in the metabolic rate per gram of tissue [1]. By multiplying the metabolic rate or SUV of a lesion by its volume to obtain “total lesion glycolysis” or TLG, this limitation may be overcome [2]. Various methodologies have been described to define tumour volumes using FDG PET. Adaptive methods that use a variable, tumour to background dependent threshold above which voxels with an activity concentration superior to this threshold are included in the tumour volume have proven to be the most robust [3, 4]. While these adaptive methodologies apply well when confronted with untreated, large tumours with high levels of metabolic activity, accurate assessment of tumour volume and thus also of TLG can prove challenging during or after treatment. Firstly, with decreasing tumour volumes, the accuracy of SUVs is affected by partial volume effects (PVE) resulting in underestimation of the true SUV values [5–7]. PVE will also influence the determination of the tumour volume through modification of the optimal threshold value for tumour delineation. Secondly, if the contrast between the tumour and surrounding normal tissue is low, no clear tumour borders can be identified. In this case, it could be argued that when using PET/CT, tumour volumes might be derived from CT images. However, quantification of tumour volumes by means of CT has proven to be quite labour intensive. Finally, tumour volume measurements by thresholding on PET images are dependent on the scan time; volumes decrease significantly and the variability of the volume is significantly larger for images with scan times below 3 min per bed position when compared to images with scan times of  $\geq 3$  min [8]. Consequently, easily applicable and straightforward methodologies assessing TLG, though avoiding PVE and tumour volume assessment by thresholding, and that are less count dependent, are of major interest. In this paper, such a method is described. The method assesses the total excess SUV above the tumour background: the standardized added metabolic activity (SAM). Phantom data as well as initial clinical validation data are presented.

## Materials and methods

### Phantom study

A Jaszczak Deluxe phantom was filled with a background concentration of 5.92 kBq/ml (0.16  $\mu$ Ci/ml)  $^{18}$ F-FDG. The spheres of the Jaszczak phantom, having a diameter of 9.5, 12.7, 15.9, 19.1, 25.4 and 31.8 mm, were filled with a 5.4 times higher concentration. Data acquisition was performed during 10 min with a Philips Allegro PET system in 2 mm<sup>3</sup> voxel size. Subsequently, additional activity was added to the background of the phantom to obtain an activity ratio of

3.64. Then a new data set was collected for 10 min. This procedure was repeated a third time using a ratio for spheres to background of 2.0.

Images were reconstructed using standard, commercially available reconstruction software (Philips). For the determination of SAM, two volumes of interest (VOIs) were delineated using PMOD software (version 3.1, PMOD Technologies Ltd., Zurich, Switzerland). The first VOI (VOI<sub>1</sub>) was delineated around the sphere and a second, larger VOI (VOI<sub>2</sub>) around VOI<sub>1</sub>. In all cases, VOI<sub>1</sub> was placed at a sufficient distance from the sphere borders in order to avoid PVE and to ensure that no spillover from the sphere to VOI<sub>2</sub> occurred.

The exact volumes of the spheres were determined by weighing the filled as well as the empty spheres. The measured inside volumes were 18.3, 11.4, 5.68, 2.13, 1.19 and 0.58 cm<sup>3</sup>. The respective volumes of the plastic shells were 12.3, 6.53, 4.13, 2.25, 1.60 and 1.36 cm<sup>3</sup>. These volumes were subtracted from VOI<sub>1</sub> for each sphere since they contain no activity (volume VOI<sub>1</sub> corrected). Without this adjustment the background correction would be overcompensating and more importantly, it would influence the result for each sphere differently. Errors would become especially large in the smallest spheres where the volume of the plastic becomes comparable to or even larger than the inner sphere volumes. Based on the volume and total SUV count of VOI<sub>1</sub> (corrected for the sphere shell volume) and VOI<sub>2</sub>, the SAM of the spheres were derived as follows:

First, mean background activity (BG in SUV) was derived by the formula:

$$\text{Mean BG} = (\text{Total SUV VOI}_2 - \text{Total SUV VOI}_1) / (\text{Volume of VOI}_2 - \text{Volume of VOI}_1)$$

in which “total SUV” is the multiplication of the mean SUV by the respective volume. Subsequently, SAM was calculated using the following formula:

$$\text{SAM} = \text{Total SUV VOI}_1 - (\text{mean BG} * \text{Volume VOI}_{1\text{corrected}})$$

Finally, normalized SAM was calculated as follows:

$$\text{norm SAM} = \text{SAM} / \text{mean BG}.$$

Measured SAM as well as normalized SAM were expressed as a percentage of the expected values which is equal to the actual activity ratio minus 1.

### Patient studies

#### Clinical settings

*Impact of acquisition time on SAM* Fifteen patients suffering from primary squamous cell carcinoma (PSCC) of the

head and neck were included in this substudy. All patients underwent a whole-body FDG PET/CT examination and subsequently a 10-min list-mode acquisition PET/CT study of the head and neck region. From the list-mode data, image frames with scan times of 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10 min were reconstructed. On the reconstructed images, SAM of the primary tumour as well as  $SUV_{max}$  values were determined. The effects of the scan time on SAM and  $SUV_{max}$  values of the primary tumour were assessed. SAM values derived from images with a duration of 10 min ( $SAM_{10}$ ) were considered the gold standard. SAM values of the remaining nine images (duration 1–9 min) were expressed as a percentage of the 10-min SAM value. The same measurements were performed for normalized SAM and  $SUV_{max}$ .

**SAM for response assessment** In this substudy, 19 patients (11 women) were included. All patients suffered from colorectal carcinoma metastasized to the liver for which they were treated by means of chemotherapy (FOLFOX, FOLFIRI). All patients underwent a PET/CT examination prior to treatment initiation (scan 1) and after 3–6 cycles of chemotherapy (scan 2). SAM values of the liver lesions on scan 1 ( $SAM_1$ ) as well as  $SUV_{max}$  values were determined and the percentage reduction between both was calculated as follows:

$$\Delta SAM = (SAM_1 - SAM_2) / SAM_1 \times 100$$

$$\Delta SUV_{max} = (SUV_{max1} - SUV_{max2}) / SUV_{max1} \times 100, \text{ were obtained.}$$

The complete disappearance of abnormal FDG uptake was considered a 100 % reduction. Results obtained were related to Response Evaluation Criteria in Solid Tumors (RECIST) response outcome [9]. Receiver-operating characteristic (ROC) curve analysis was performed to assess the discriminatory power of  $\Delta SAM$  and  $\Delta SUV_{max}$  to separate responders from non-responders as defined by RECIST response outcome.

#### Data acquisition

All patients were injected with a dose of 3.7 MBq (100  $\mu$ Ci) of  $^{18}F$ -FDG per kg body weight. PET/CT acquisition was started between 45 and 60 min after the tracer administration with a duration of 3 min per bed position (matrix size 4 mm<sup>3</sup> voxel). Images were acquired from the base of the skull to the proximal third of the femora.

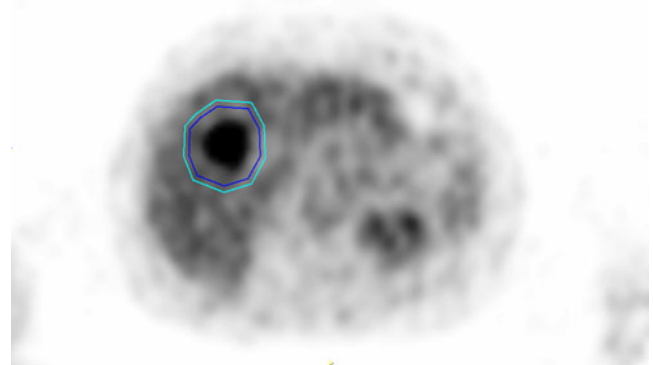
The Gemini PET/CT imaging system (Philips Co., Cleveland, OH, USA) consists of a gadolinium oxyorthosilicate (GSO) full-ring PET scanner with 5.0 mm spatial resolution and a 16-slice helical CT scanner. Standard software recommended by the manufacturer was used for acquisition and processing. Reconstruction was performed

using an iterative 3-D row action maximum likelihood algorithm (RAMLA), with 2 iterations and using generalized Kaiser-Bessel functions (blobs). Scatter and attenuation correction were applied using the single scatter simulation algorithm and measured attenuation correction, respectively.

Patients in substudy 1 subsequently underwent an additional PET/CT study in list-mode acquisition of 10 min duration. First a low-dose CT [30 mA, 120 kV, field of view (FOV) 600 mm, 0.5 s rotation time, pitch 0.9, collimation 16×1.5 mm] was performed with the arms at the side of the torso. Immediately after the CT acquisition a list-mode acquisition with a duration of 10 min was performed. From the list-mode data, image frames of 1–10 min duration were reconstructed and corrected for scatter and attenuation.

#### Image analysis (calculation of SAM)

A volume of interest ( $VOI_1$ ) was drawn using PMOD software over the primary tumour lesions of patients of substudy 1 on all reconstructed image sets (1- to 10-min reconstructed images) as well as over all individual metastatic liver lesions of patients of substudy 2 (scan 1 and scan 2). A second  $VOI$  ( $VOI_2$ ) was delineated around  $VOI_1$  directed to a small zone of homogeneous background. The borders of  $VOI_1$  were set at a reasonable distance from the tumour lesions in order to avoid PVE and to ensure that no spillover from the tumour lesion to  $VOI_2$  occurred (see Fig. 1). SAM and normalized SAM values were derived using the formulas described in the phantom study, but without the volume correction for  $VOI_1$ , as this does not apply on patients. In patients with multiple liver metastases (substudy 2) SAM was calculated as the sum of the individual SAMs of the lesions.



**Fig. 1**  $^{18}F$ -FDG PET scan of a patient with colorectal carcinoma metastasized to the liver. A volume of interest ( $VOI_1$ , blue  $VOI$ ) was set at a reasonable distance from the metastatic lesion in order to avoid PVE. A second  $VOI$  ( $VOI_2$ , green  $VOI$ ) was delineated around  $VOI_1$ . Subtracting the background activity in  $VOI_2 - VOI_1$  from  $VOI_1$  yields SAM

## Results

### Phantom study

Table 1 shows that for a ratio of 5.4, all of the normalized SAM values/sphere volume remain constant for all spheres, with a mean value of 91.9 % (SD 6.8 %) of the expected value of 4.4. For the phantom study with a ratio of 3.64, the 0.58 ml sphere was no longer visible. For the five remaining spheres a mean value of 97.6 % (SD 3.6 %) of the expected value of 2.64 was obtained. For the lowest ratio of 2.0, four spheres were still visible. The smallest one had a volume of 2.1 ml. A mean value of 96.0 % (SD 6.2 %) of the expected 1.0 was calculated. Overall for the 3 acquisitions and the 15 visible spheres, the mean normalized SAM/sphere volume calculated was 94.9 % (SD 5.9 %) of the expected value. These results show that SAM per unit of volume remains unchanged for smaller spheres, proving that this method is not influenced by PVE and can be used for small lesions. Since in patients the tumour size is not known, SAM/tumour volume cannot be calculated. Instead, SAM or SAM normalized for background activity can be used.

### Clinical studies

*Impact of acquisition time on SAM* Absolute mean (+ range) SAM<sub>10</sub> values, mean BG<sub>10</sub> values and normalized SAM<sub>10</sub> values of PSCC were 70.9 (6.9–105.4), 0.7 (0.4–1.1) and 73.8 (13.3–255.6), respectively. Absolute mean (+ range) SUV<sub>max</sub> values of PSCC were 5.8 (2.6–10.5).

SAM values, normalized SAM values and SUV<sub>max</sub> values of PSCC expressed as a percentage of the 10-min value for the different acquisition durations (1–10 min) are shown in Table 2. While SAM decreased significantly with shorter scan time (Friedman,  $p < 0.0001$ ), the differences proved however small. The mean difference between the 1-min

acquisition and the 10-min reference was only 4 % (SD 5 %; range 88–105 %). This decrease with shorter scan time was not seen for normalized SAM, but the dividing operation increased the statistical noise resulting in a larger variability (SD 10 %; range 88–121 %).

SUV<sub>max</sub> was not significantly different between images with various scan times (Friedman,  $p = 0.2$ ). The variability of SUV<sub>max</sub> values proved however slightly larger than that of SAM values as illustrated by the SD of 6.4 % (range 86–110 %).

*SAM for response assessment* No systematic differences in starting times (between baseline and response scans, between responders and non-responders or a combination of both) were observed ( $p \geq 0.2$ ).

SAM1 values of liver lesions ranged from 18.3 to 2,111 (mean 196; SD 487). SUV<sub>max</sub>1 values ranged from 1.7 to 14.6 (mean 5.3; SD 3.0).

According to the RECIST criteria as applied on CT we identified 8 responders and 11 non-responders. SAM1 and  $\Delta$ SAM values for responders versus non-responders were 57 (SD 119) versus 297 (SD 625) for SAM1 ( $p = 0.2$ ) and 99 % (SD 3 %) versus 32 % (SD 44 %) for  $\Delta$ SAM ( $p = 0.001$ ), respectively. The area under the ROC curve (AUC) for SAM1 and  $\Delta$ SAM were 0.74 and 1.0, respectively, suggesting that  $\Delta$ SAM may accurately assess response to treatment of liver metastases in patients suffering from colorectal cancer (CRC) (see Fig. 2). Using a cutoff of 85 % reduction for  $\Delta$ SAM, responders could be separated from non-responders with a sensitivity of 100 % and a specificity of 100 %.

SUV<sub>max</sub>1 and  $\Delta$ SUV<sub>max</sub> values in responders versus non-responders were 3.9 (SD 2.4) versus 6.3 (SD 3.1) for SUV<sub>max</sub>1 ( $p = 0.08$ ) and 94 % (SD 17 %) versus 7 % (SD 40 %) for  $\Delta$ SUV<sub>max</sub> ( $p = 0.0001$ ), respectively. The AUC for SUV<sub>max</sub>1 and  $\Delta$ SUV<sub>max</sub> were 0.79 and 0.99, respectively, suggesting that  $\Delta$ SUV<sub>max</sub> may accurately assess response to treatment of liver metastases in patients suffering from CRC (see Fig. 2). Using a cutoff of 48 % reduction of  $\Delta$ SUV<sub>max</sub>, responders could be separated from non-responders with a sensitivity of 100 % and a specificity of 91 %.

The AUC of  $\Delta$ SAM and  $\Delta$ SUV<sub>max</sub> proved to be not significantly different ( $p = 0.6$ ).

## Discussion

The main problem for accurate quantification in PET is the relatively low resolution of the PET camera. Both quantification of uptake values and volume determination are affected by the PVE [6].

In this study, using a Jaszczak phantom with a homogeneous background in order to simulate normal liver activity, we show

**Table 1** Normalized SAM/sphere volume

|                   | Sphere volume (ml) |       |      |       |              |              |
|-------------------|--------------------|-------|------|-------|--------------|--------------|
|                   | 18.3               | 11.4  | 5.7  | 2.1   | 1.19         | 0.58         |
| Ratio=5.40        |                    |       |      |       |              |              |
| Expected SAM=4.40 | 4.05               | 4.14  | 3.68 | 4.13  | 3.75         | 4.51         |
| % of expected     | 92.2               | 94.1  | 83.6 | 93.8  | 85.2         | 02.5         |
| Ratio=3.64        |                    |       |      |       |              |              |
| Expected SAM=2.64 | 2.50               | 2.74  | 2.58 | 2.55  | 2.52         | <sup>a</sup> |
| % of expected     | 94.6               | 103.8 | 97.7 | 96.6  | 95.5         | <sup>a</sup> |
| Ratio=2.00        |                    |       |      |       |              |              |
| Expected SAM=1.00 | 0.94               | 0.89  | 0.97 | 1.04  | <sup>a</sup> | <sup>a</sup> |
| % of expected     | 94.0               | 89.0  | 97.0 | 104.0 | <sup>a</sup> | <sup>a</sup> |

<sup>a</sup> Spheres no longer visible

**Table 2** SAM, normalized SAM and  $SUV_{max}$  expressed as a percentage of the 10-min value

|        | SAM   |     |       |       | SAM normalized |      |       |       | $SUV_{max}$ |     |       |       |
|--------|-------|-----|-------|-------|----------------|------|-------|-------|-------------|-----|-------|-------|
|        | Mean  | SD  | Min.  | Max.  | Mean           | SD   | Min.  | Max.  | Mean        | SD  | Min.  | Max.  |
| 1 min  | 96.0  | 5.0 | 88.3  | 105.1 | 99.9           | 10.0 | 88.0  | 121.3 | 99.6        | 6.4 | 85.7  | 109.5 |
| 2 min  | 97.1  | 2.7 | 90.2  | 100.8 | 98.1           | 5.7  | 87.4  | 108.5 | 98.4        | 3.9 | 93.5  | 108.4 |
| 3 min  | 98.1  | 2.6 | 92.5  | 102.5 | 98.1           | 4.3  | 90.6  | 108.6 | 99.7        | 2.5 | 95.7  | 105.5 |
| 4 min  | 99.0  | 2.0 | 95.0  | 103.9 | 99.0           | 2.6  | 94.4  | 104.0 | 100.0       | 2.0 | 96.7  | 103.1 |
| 5 min  | 99.0  | 1.7 | 96.1  | 102.8 | 98.6           | 1.9  | 96.0  | 101.8 | 99.1        | 2.0 | 95.7  | 102.4 |
| 6 min  | 99.1  | 1.4 | 96.8  | 101.6 | 98.9           | 1.7  | 96.8  | 102.5 | 99.2        | 1.9 | 96.1  | 103.2 |
| 7 min  | 99.4  | 1.0 | 97.9  | 101.5 | 99.5           | 1.2  | 98.1  | 101.4 | 99.1        | 1.8 | 96.8  | 104.2 |
| 8 min  | 99.5  | 0.9 | 97.3  | 100.8 | 99.5           | 1.1  | 96.7  | 100.8 | 99.3        | 1.2 | 97.2  | 102.4 |
| 9 min  | 99.8  | 0.6 | 98.3  | 100.7 | 99.8           | 0.9  | 97.7  | 102.0 | 99.8        | 1.0 | 98.1  | 101.7 |
| 10 min | 100.0 | 0.0 | 100.0 | 100.0 | 100.0          | 0.0  | 100.0 | 100.0 | 100.0       | 0.0 | 100.0 | 100.0 |

SD standard deviation

that SAM is sphere size independent, confirming that, when using SAM as a quantitative parameter, problems related to PVE and threshold values for tumour delineation are avoided.

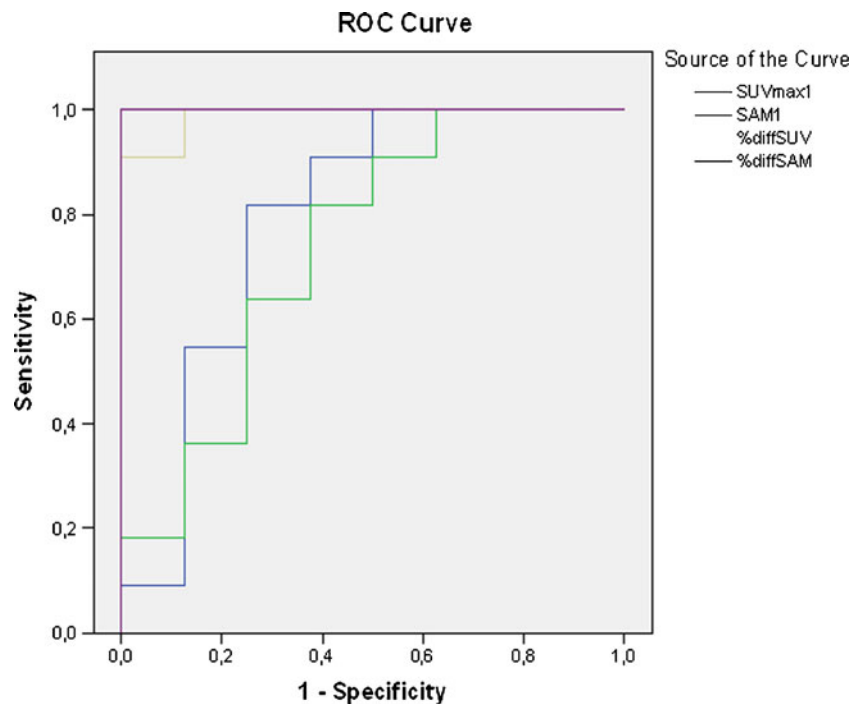
Still there are some factors that might affect the accuracy and precision of SAM such as the calibration of the PET scanner, injected dose and patient weight. For this reason, another parameter, normalized SAM (SAM/mean BG) was introduced. As the aforementioned sources of error affect both SAM and the background similarly, normalized SAM is no longer affected by them. However, the dividing operation increases the statistical noise.

Another factor affecting the accuracy of all quantitative parameters is the reconstruction algorithm. Despite the fact

that we found consistent SAM values for different sphere volumes, the values proved to be systematically lower than expected. It is well known that the reconstruction algorithm and the adjustment of its parameters can influence the quantification accuracy [10]. Also, inadequate scatter correction might be the cause of these consistently lower values. Especially in large axial FOV 3-D PET scanners the contribution of scatter from outside the FOV is a challenging issue that needs to be addressed [11].

Tumour volume assessment on FDG PET imaging is gaining wide clinical interest outside the field of radiotherapy planning, where its primary interest is biological target volume delineation [3, 12, 13].

**Fig. 2** ROC curves for  $SUV_{max1}$ , SAM1,  $\Delta SUV_{max}$  and  $\Delta SAM$



First, in line with studies performing volumetric tumour staging by means of CT or MRI [14–16], it was recently shown that the metabolic tumour volume as assessed by FDG PET contains independent prognostic information on disease-free and overall survival in various types of cancer [17–22]. The assessment of the metabolic volume as described in these reports was performed using either manual or automated delineation of the tumour, e.g. region-growing or segmentation-based methods. Automated tumour contouring or delineation is more robust when compared to manual delineation. Unfortunately, in clinical practice the less reproducible manual method is most often the method used as most tumours present with an irregular pattern of FDG uptake and a corresponding complex tumour demarcation. More recently, some of the more advanced automated methods, e.g. gradient-based methods, the improved fuzzy C-means algorithm and the fuzzy locally adaptive Bayesian algorithm, have proved to be more robust and able to deal efficiently with complex tumour demarcation on PET images acquired on different PET/CT systems, using different voxel sizes and different tracers, and this both in phantom models as well as in patient studies [3, 4, 23, 24]. Thus, these segmentation algorithms may provide a robust and repeatable tool to aid physicians in determining functional volumes in PET imaging.

Opposed to automated techniques, SAM avoids the problem of tumour delineation as it assumes that the tumour tissue develops on top of the normal tissue, hence the subtraction of local surrounding normal tissue activity. Furthermore, SAM is only minimally influenced by the acquisition duration. In patients suffering from PSCC of the head and neck, the mean difference between SAM measurements derived from 1 min versus 10 min acquisition was only 4 %; for volume measurements using automated thresholding a mean difference in a similar setup of 16 % was reported previously [8]. While  $SUV_{max}$  values proved to be not significantly different between images of variable scan time, the variability was slightly larger (SD 6.4 % for  $SUV_{max}$  versus 5 % for SAM).  $SUV_{max}$  is by definition obtained from the hottest voxel and thus more prone to be affected by noise when compared to SAM [25].

It might be argued that by adopting SAM the true tumour volume is underestimated. However, the concept of SAM and background subtraction is fully in line with the current clinical practice in FDG PET interpretation in which normal scan findings or a complete response are defined as the absence of detectable lesions against the normal background.

Recently, the PET Response Criteria in Solid Tumors (PERCIST) framework was introduced for tumour response evaluation, in which the authors also identify TLG as an attractive parameter. PERCIST suggests that for tumour volume delineation a threshold should be used equal to 2-3 SDs above the mean SUV in healthy liver parenchyma of

the patient [26]. Similarly, by subtracting background activity, SAM takes into account the variability in metabolic activity of normal tissue between scans, but SAM also takes it one step further as the background SUV is tissue specific.

The basic principle behind PERCIST is that tumour response is a continuous and time-dependent variable which is expressed as a percentage change in  $SUV_{peak}$  (average activity within a spherical region of interest measuring 1.2 cm in diameter centred at the most active portion of the tumour), rather than  $SUV_{max}$  between the pre- and post-treatment scans. PERCIST further specifies that the  $SUV_{peak}$  is to be obtained on the single most active lesion on each scan, which may be located in a different lesion on a follow-up scan. It is further recommended that a sum of  $SUV_{peak}$  of up to five target lesions (no more than two per organ) would be measured as a secondary determinant of response. In those lesions that are smaller than 2 cm in diameter, assessment of TLG and  $SUV_{peak}$  activity will be underestimated due to the PVE [6, 23]. Furthermore, the reduction in size of lesions <2 cm in diameter and the reduction in size of lesions >2 cm to a size prone to PVE will result in an overestimation of the metabolic response. Accordingly, the spread on the range of percentage change in TLG and  $SUV_{peak}$  in partial responders may vary significantly depending on the size of the lesions included for PERCIST analysis. As SAM is not influenced by the PVE, hypothetically, it is likely to significantly decrease the range of percentage change in partial responders. In the patient series presented, only one RECIST responder presented with a partial response on PET; as such the aforementioned hypothesis remains to be proven. As nearly all responders were complete responders, the AUC of 1 is likely to decrease in a setting which includes more partial responders. Nevertheless, the findings presented suggest that SAM which incorporates both the primary and secondary determinants of response suggested by PERCIST is a promising parameter for tumour response assessment of liver metastases. Additional studies in larger patient populations and different tumour pathologies, including tumours with more heterogeneous background activity, as well as in a mix of tumour responses are warranted.

#### Study limitations

Because of the retrospective character of the clinical study included, FDG PET/CT images were acquired within a range of 45-60 min following injection of FDG. However, no systematic differences in starting times (between baseline and response scans, between responders and non-responders or a combination of both) were observed.

The presented methodology has only been tested in conditions with a homogeneous background, namely the Jaszczak phantom and liver parenchyma. As already mentioned above, the performance of the presented methodology in other

tumour pathologies with more heterogeneous background is not known. Phantom and clinical studies in this regard should also be performed.

In the clinical substudy, the change in SAM of lesions is related to the RECIST response. The use of these morphometric criteria is based on the abundant evidence showing that agents that produce tumour shrinkage will most likely also prolong survival [9]. In the last decade, PET has proven to be a useful tool in response assessment, even outperforming morphometric evaluation in several indications [26–28]. Nevertheless, the RECIST criteria are still considered the gold standard in assessing treatment response by imaging. Obviously, further research is warranted to identify a possible relationship between SAM and progression-free or overall survival.

Another potential limitation has to be addressed. While SAM assesses the added metabolic activity on top of the background activity, it does not provide information on tumour volume or mean tumour SUV. In some applications, these two separate parameters are of interest. Accordingly, SAM should be considered as a complementary parameter that provides additional information to more commonly used PET response parameters like SUV values or metabolic tumour volume.

## Conclusion

Similarly to TLG, the SAM method takes into account the tracer uptake as well as the size of the lesion, though avoiding the need of thresholding or another delineation method. The phantom study has demonstrated that the presented method does not suffer from PVE. Preliminary clinical data show that SAM has a potential role in tumour response assessment by means of FDG PET of liver metastases, though additional clinical studies in larger patient populations and different tumour pathologies, including tumours with more heterogeneous background activity, as well as in a mix of tumour responses are warranted.

**Acknowledgments** No financial support was received. The hours of practical and administrative work by Dr. Berwouts Dieter and Mw. Blanken Tamara as well as the support of the whole administrative, nursing, radiopharmaceutical, engineering and medical staff of the nuclear medicine department at the University Hospital of Ghent were all much appreciated.

## References

- Weber W, Figlin R. Monitoring cancer treatment with PET/CT: does it make a difference? *J Nucl Med* 2007;48:36S–44S.
- Larson S, Erdi Y, Akhurst T, Mazumdar M, Macapinlac HA, Finn RD, et al. Tumor treatment response based on visual and quantitative changes in global tumor glycolysis using PET-FDG imaging. The visual response score and the change in total lesion glycolysis. *Clin Positron Imaging* 1999;2:159–71.
- Zaidi H, El Naqa IE. PET-guided delineation of radiation therapy treatment volumes: a survey of image segmentation techniques. *Eur J Nucl Med Mol Imaging* 2010;37:2165–87.
- Tylski P, Stute S, Grotus N, Doyeux K, Hapdey S, Gardin I, et al. Comparative assessment of methods for estimating tumor volume and standardized uptake value in (18)F-FDG PET. *J Nucl Med* 2010;51:268–76.
- Weber WA. Quantitative analysis of PET studies. *Radiother Oncol* 2010;96:308–10.
- Soret M, Bacharach S, Buvat I. Partial-volume effect in PET tumor imaging. *J Nucl Med* 2007;48:932–45.
- Hoetjes N, van Velden F, Hoekstra O, Krak NC, Lammertsma AA, Boellaard R. Partial volume correction strategies for quantitative FDG PET in oncology. *Eur J Nucl Med Mol Imaging* 2010;37:1679–87.
- Goethals I, D'Asseler Y, Dobbeleir A, Deblaere K, Ham H. The effect of acquisition time on visual and semi-quantitative analysis of F-18 FDG-PET studies in patients with head and neck cancer. *Nucl Med Commun* 2010;31:227–31.
- Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* 2009;45:228–47.
- Sattler B, Lee J, Lonsdale M, Coche E. PET/CT (and CT) instrumentation, image reconstruction and data transfer for radiotherapy planning. *Radiother Oncol* 2010;96:288–97.
- Zaidi H. Comparative evaluation of scatter correction techniques in 3D positron emission tomography. *Eur J Nucl Med* 2000;27:1813–26.
- Lammering G, De Ruyscher D, van Baardwijk A, Baumert BG, Borger J, Lutgens L, et al. The use of FDG-PET to target tumors by radiotherapy. *Strahlenther Onkol* 2010;186:471–81.
- Bussink J, Kaanders J, van der Graaf W, Oyen W. PET-CT for radiotherapy treatment planning and response monitoring in solid tumors. *Nat Rev Clin Oncol* 2011;8:233–42.
- Doweck I, Denys D, Robbins KT. Tumor volume predicts outcome for advanced head and neck cancer treated with targeted chemoradiotherapy. *Laryngoscope* 2002;112:1742–9.
- Knegiens J, Hauptmann M, Pameijer F, Balm AJ, Hoebbers FJ, de Bois JA, et al. Tumor volume as prognostic factor in chemoradiation for advanced head and neck cancer. *Head Neck* 2011;33:375–82.
- Plataniotis G, Theofanopoulou M, Kalogera-Fountzila A, Haritanti A, Ciuleanu E, Ghilezan N, et al. Prognostic impact of tumor volumetry in patients with locally advanced head-and-neck carcinoma (non-nasopharyngeal) treated by radiotherapy alone or combined radiochemotherapy in a randomized trial. *Int J Radiat Oncol Biol Phys* 2004;59:1018–26.
- Yan H, Wang R, Zhao F, Zhu K, Jiang S, Zhao W, et al. Measurement of tumor volume by PET to evaluate prognosis in patients with advanced non-small cell lung cancer treated by non-surgical therapy. *Acta Radiol* 2011;52:646–50.
- Gulec S, Suthar R, Barot T, Pennington K. The prognostic value of functional tumor volume and total lesion glycolysis in patients with colorectal cancer liver metastases undergoing (90)Y selective internal radiation therapy plus chemotherapy. *Eur J Nucl Med Mol Imaging* 2011;38:1289–95.
- Chung H, Kim J, Han K, Eo JS, Kang KW, Park NH, et al. Prognostic value of metabolic tumor volume measured by FDG-PET/CT in patients with cervical cancer. *Gynecol Oncol* 2011;120:270–4.
- Schinagl D, Span P, Oyen W, Kaanders J. Can FDG PET predict radiation treatment outcome in head and neck cancer? Results of a prospective study. *Eur J Nucl Med Mol Imaging* 2011;38:1449–58.

21. Hatt M, Visvikis D, Pradier O, Cheze-le-Rest C. Baseline 18F-FDG PET image-derived parameters for therapy response prediction in oesophageal cancer. *Eur J Nucl Med Mol Imaging* 2011;38:1595–606.
22. Cazaentre T, Morschhauser F, Vermandel M, Betrouni N, Prangère T, Steinling M, Huglo D. Pre-therapy 18F-FDG PET quantitative parameters help in predicting the response to radioimmunotherapy in non-Hodgkin lymphoma. *Eur J Nucl Med Mol Imaging* 2010;37:494–504.
23. Hatt M, Cheze Le Rest C, Albaraghach N, Pradier O, Visvikis D. PET functional volume delineation: a robustness and repeatability study. *Eur J Nucl Med Mol Imaging* 2011;38:663–72.
24. Hatt M, Cheze-Le Rest C, Aboagye E, Kenny LM, Rosso L, Turkheimer FE, et al. Reproducibility of 18F-FDG and 3'-deoxy-3'-18F-fluorothymidine PET tumor volume measurements. *J Nucl Med* 2010;51:1368–76.
25. Boellaard R, Krak N, Hoekstra O, Lammertsma A. Effects of noise, image resolution, and ROI definition on the accuracy of standard uptake values: a simulation study. *J Nucl Med* 2004;45:1519–27.
26. Wahl R, Jacence H, Kasamon Y, Lodge MA. From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors. *J Nucl Med* 2009;50:122S–50S.
27. Juweid ME, Cheson BD. Positron-emission tomography and assessment of cancer therapy. *N Engl J Med* 2006;354:496–507.
28. Weber WA, Wieder H. Monitoring chemotherapy and radiotherapy of solid tumors. *Eur J Nucl Med Mol Imaging* 2006;33:27–37.