ORIGINAL ARTICLE

Evaluation of a cumulative SUV-volume histogram method for parameterizing heterogeneous intratumoural FDG uptake in non-small cell lung cancer PET studies

**Floris H. P. van Velden · Patsuree Cheebsumon ·
Maqsood Yaqub · Egbert F. Smit · Otto S. Hoekstra ·
Adriaan A. Lammertsma · Ronald Boellaard**

Abstract

Purpose Standardized uptake values (SUV) are commonly used for quantification of whole-body [18F]fluoro-2-deoxy-D-glucose (FDG) positron emission tomography (PET) studies. Changes in SUV following therapy, however, only provide a proper measure of response in case of homogeneous FDG uptake in the tumour. The purpose of this study was therefore to implement and characterize a method that enables quantification of heterogeneity in tumour FDG uptake.

Methods Cumulative SUV-volume histograms (CSH), describing % of total tumour volume above % threshold of maximum SUV (SUV_{max}), were calculated. The area under a CSH curve (AUC) is a quantitative index of tumour uptake heterogeneity, with lower AUC corresponding to higher degrees of heterogeneity. Simulations of homogeneous and heterogeneous responses were performed to assess the value of AUC-CSH for measuring uptake and/or response heterogeneity. In addition, partial volume correction and image denoising was applied prior to calculating AUC-CSH. Finally, the method was applied to a number of human FDG scans.

Results Partial volume correction and noise reduction improved CSH curves. Both simulations and clinical examples showed that AUC-CSH values corresponded with level of tumour heterogeneity and/or heterogeneity in response. In contrast, this correspondence was not seen with SUV_{max} alone. The results indicate that the main advantage of AUC-CSH above other measures, such as 1/COV (coefficient of variation), is the possibility to measure or normalize AUC-CSH in different ways.

Conclusion AUC-CSH might be used as a quantitative index of heterogeneity in tracer uptake. In response monitoring studies it can be used to address heterogeneity in response.

Keywords Positron emission tomography (PET) · Standardized uptake value (SUV) · Intratumoural heterogeneity · Cumulative SUV-volume histogram (CSH) · Intensity-volume histograms (IVH)

F. H. P. van Velden (✉) · P. Cheebsumon · M. Yaqub ·
O. S. Hoekstra · A. A. Lammertsma · R. Boellaard
Department of Nuclear Medicine & PET Research,
VU University Medical Center,
PO Box 7057, 1007 MB Amsterdam, The Netherlands
e-mail: f.vvelden@vumc.nl

E. F. Smit
Department of Pulmonary Diseases,
VU University Medical Center,
Amsterdam, The Netherlands

Introduction

In non-small cell lung cancer (NSCLC), positron emission tomography (PET) has been shown to be a valuable tool not only for detecting and staging the disease, but also for response monitoring, prediction of prognosis and estimation of target volume for radiotherapy purposes [1, 2]. To date, [18F]fluoro-2-deoxy-D-glucose (FDG) is the most widely used tracer for oncological applications. Especially for response monitoring purposes, it is likely that quantitative assessment of FDG uptake will become the standard. In general, however, uptake of this tracer is not homogeneously distributed across the tumour. Factors that may contribute to intratumoural heterogeneity in FDG uptake are necrosis [3], cellular proliferation [4], blood flow [5], microvessel density [6] and hypoxia [7–9]. The standardized uptake value (SUV)

is most commonly used for (semi-)quantification of whole-body FDG PET studies [10]. A high maximum SUV (SUV_{max}) has been shown to relate to hypoxia and poor overall survival [11]. However, changes in SUV to assess response to therapy only provide a proper measure of response if there is a global change in tracer uptake, i.e. in the absence of a spatially heterogeneous response. Therefore, intratumoural FDG heterogeneity may complicate accurate response assessment with PET. As uptake of a tracer usually is not homogeneously distributed across the tumour, it is of interest to quantify heterogeneity in tumour FDG uptake before, during and after treatment.

Firstly, the distribution of FDG uptake within a tumour could provide useful information for radiation therapy treatment planning, as it would allow for specific targeting of certain areas within the tumour. Secondly, it may provide additional information when monitoring response, as it would enable identification of a mixed response within a single tumour. Finally, differences between CT anatomical volumes and PET metabolic volumes could be characterized. In current clinical practice, however, there is no simple method for quantification of intratumoural heterogeneity in FDG uptake [8].

Only a few methods have been proposed to quantify intratumoural heterogeneity in FDG uptake. O'Sullivan et al. [12, 13] proposed a method for patients with sarcoma that compares the intratumoural FDG uptake distribution with an elliptic solid object with homogeneous density. Recently, intensity-volume histograms (IVH) or cumulative SUV-volume histograms (CSH) have been proposed by El Naqa et al. [14] as a novel way to characterize heterogeneity in intratumoural tracer uptake. These histograms are similar to dose-volume histograms frequently used in radiotherapy [15]. In CSH the per cent volume of a tumour (derived from CT or from PET-based (semi-)automatic tumour delineation methods [16]) with an SUV above a certain threshold is plotted against that threshold value, which is varied from 0 to 100% of SUV_{max}. The area under the CSH (AUC-CSH) may be a quantitative index of tracer uptake heterogeneity and/or heterogeneous tumour response [17]. Any method to characterize heterogeneity, however, will treat both partial volume effects and noise as heterogeneity and therefore partial volume correction and image denoising must be applied prior to calculating AUC-CSH [18].

The purpose of this study was to evaluate to which extent CSH provides additional information about tumour response (and its heterogeneity) over SUV_{max} alone. In addition, the impact of partial volume effects and noise were evaluated by applying partial volume correction and image denoising prior to calculating AUC-CSH. To this end, lung cancer simulations were performed and the method was applied to several FDG PET studies of tumours with variable heterogeneity.

Materials and methods

Cumulative SUV-volume histograms

Four strategies for calculating CSH were investigated. As mentioned above, CSH is normally obtained by plotting the per cent volume of a tumour with an SUV above a certain threshold against that threshold, which is varied from 0 to 100% of SUV_{max}. The AUC of this plot (AUC-CSH) is a quantitative index of uptake heterogeneity, where lower values correspond with increased heterogeneity. However, in this case AUC-CSH is independent of SUV_{max} and volume, and therefore it may not be useful for response monitoring, where changes in SUV_{max} or metabolic volume should also be taken into account. To do so, three modifications of CSH were defined, taking into account changes in metabolic volume (CSH^V), SUV_{max} (CSH^S) or both (CSH^{SV}). This was achieved by not plotting relative (percentage) SUV_{max} or metabolic volume data, but rather absolute values or values relative to the baseline SUV_{max} or volume. Normalized AUC-CSH values (relative to the baseline) are calculated by:

$$normalized\ AUC - CSH = \frac{AUC - CSH^{response}}{AUC - CSH^{baseline}}$$

$$normalized\ AUC - CSH^V = \frac{AUC - CSH^{response}}{AUC - CSH^{baseline}} \times \frac{volume^{response}}{volume^{baseline}}$$

$$normalized\ AUC - CSH^S = \frac{AUC - CSH^{response}}{AUC - CSH^{baseline}} \times \frac{SUV_{max}^{response}}{SUV_{max}^{baseline}}$$

$$normalized\ AUC - CSH^{SV} = \frac{AUC - CSH^{response}}{AUC - CSH^{baseline}} \times \frac{SUV_{max}^{response}}{SUV_{max}^{baseline}} \times \frac{volume^{response}}{volume^{baseline}}$$

All four strategies were evaluated with and without partial volume correction and with and without image denoising. Partial volume correction was performed using image-based Van Cittert deconvolution (CIT) [19, 20]. During each iterative deconvolution step (i.e. iteration, five in total) the update of the image was penalized using a Gibbs prior (weight of 0.25) applying a neighbourhood of one voxel in all directions [21]. The Gibbs prior calculates the summed difference between

Table 1 Description of the various simulated tumour responses for simulation 1

Response tumour type	Code	Description of response
No response	NR1	Tumour diameter remained at 8 cm with SUV remaining at 6
Homogeneous	HO1	Tumour diameter remained at 8 cm with SUV decreasing to 3
	HO2	Tumour diameter remained at 8 cm with SUV increasing to 12
	HO3	Tumour diameter decreased to 4 cm with SUV remaining at 6
	HO4	Tumour diameter decreased to 4 cm with SUV decreasing to 3
	HO5	Tumour diameter decreased to 4 cm with SUV increasing to 12
Heterogeneous	HE1	Tumour diameter remained at 8 cm with SUV remaining at 6 (outer rim), having an inner core of 4 cm diameter with an SUV of 3
	HE2	Tumour diameter remained at 8 cm with SUV remaining at 6 (outer rim), having an inner core of 6 cm diameter with an SUV of 3
	HE3	Tumour diameter remained at 8 cm with SUV remaining at 6 (outer rim), having an inner core of 4 cm diameter with an SUV of 12
	HE4	Tumour diameter remained at 8 cm with SUV remaining at 6 (outer rim), having an inner core of 6 cm diameter with an SUV of 12

the central voxel and its direct neighbours. Inclusion of the prior reduces voxel variability in uniform areas of the image. Masks that were used to determine metabolic volumes were eroded with a single voxel in all dimensions to compensate for remaining sampling error (due to the use of voxels) at the borders of the masks. In addition, PET images were denoised using an edge-preserving bilateral filter (BF) [22].

Other parameters to characterize tracer uptake

In addition to AUC-CSH, the following parameters relating to tumour FDG uptake were also investigated: SUV_{max}, average SUV (SUV_{mean}) within a volume of interest (VOI), inverse standard deviation of SUV within a volume of interest (1/SD) and inverse coefficient of variation [1/COV, calculated as (SUV_{mean}/SD)×100%] within the metabolic (i.e. PET-based) or anatomical (i.e. CT-based) volume.

Simulations

The purpose of the simulations was to evaluate to which extent CSH provides additional information about tumour

response (and its heterogeneity) over SUV_{max} alone. In addition, the impact of partial volume effects and noise were evaluated by applying partial volume correction and image denoising prior to calculating AUC-CSH. To this end, two simulation studies were performed. As a starting point simulation 1 used a tumour with homogeneous FDG uptake, while for simulation 2 a tumour with heterogeneous FDG uptake was used. For both studies, various tumour responses showing different homogeneous and heterogeneous FDG uptake were simulated, as shown in Tables 1 and 2 and Fig. 1. All simulation software was implemented in-house using IDL (ITT Visual Information Solutions, Boulder, CO, USA). The effects of randoms and scatter events were not simulated in this study as this would require more complex (Monte Carlo-based) simulations. However, effects of attenuation and resolution were taken into account.

A 3-D mathematical thorax image and its corresponding μ-image, used for attenuation correction purposes, were derived from a dynamic FDG scan of a typical patient, as described in detail in Boellaard et al. [10]. In short, the procedure was as follows. For the simulated baseline (BL)

Table 2 Description of the various simulated tumour responses for simulation 2

Response tumour type	Code	Description of response
No response	NR2	Tumour diameter remained at 8 cm with SUV remaining at 6 (outer rim), having an inner core of 4 cm diameter with an SUV of 12
Homogeneous	HOA	Tumour diameter decreased to the inner core of 4 cm diameter with an SUV of 12
	HOB	Inner core of the tumour increased to 8 cm diameter, remaining at an SUV of 12
Heterogeneous	HEA	Tumour diameter remained at 8 cm with SUV remaining at 6 (outer rim), having an inner core of 4 cm diameter with an SUV of 3
	HEB	Overall tumour SUV decreased by a factor of 2, where tumour diameter remained at 8 cm with SUV decreasing to 3 (outer rim) and an inner core of 4 cm diameter with an SUV of 6
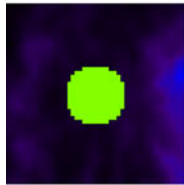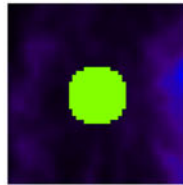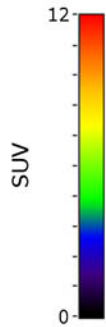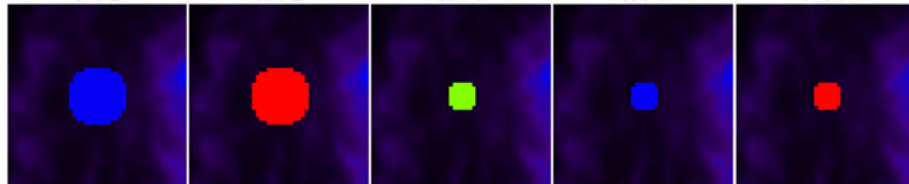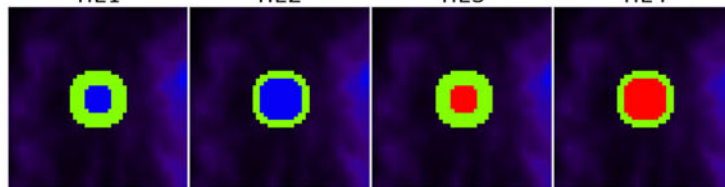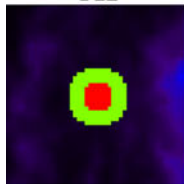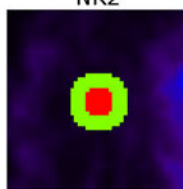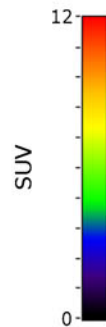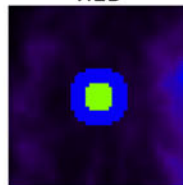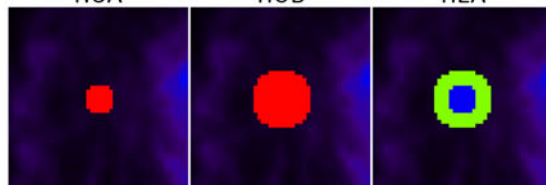
◄ **Fig. 1** Axial images of simulated mathematical baseline scans and various response scans, showing a tumour placed in the left lung. *BL* baseline, *NR* no response, *HO* homogeneous tracer uptake, *HE* heterogeneous tracer uptake. Note that these images represent ideal (i.e. free of noise and partial volume effect) scans. During simulations noise and partial volume effects were added. For simulation 1, HO and HE represent homogeneous and heterogeneous responses, respectively. However, for simulation 2, all HO and HE show heterogeneous responses, except for HEB that shows a homogeneous response

scans of simulation 1 (BL1), a tumour with homogeneous FDG uptake was placed within the lungs (SUV$_{mean}$ ~0.6) of the mathematical thorax image, using a tumour to background ratio of 10. In addition, various tumours having homogeneous and heterogeneous FDG uptake were placed within the lungs of the mathematical thorax image, as shown in Table 1 and Fig. 1. For the simulated baseline scans of simulation 2 (BL2), a tumour with heterogeneous FDG uptake was placed within the lungs of the mathematical thorax image, using a tumour (outer rim) to background ratio of 10. In addition, various tumours with homogeneous and heterogeneous FDG uptake were placed within the lungs of the mathematical thorax image, as shown in Table 2 and Fig. 1. For both simulations, these tumours were also placed in the lungs of the μ-image, having a μ-value equal to that observed in soft tissue at 511 keV (0.095 cm^{-1}). All of these

Fig. 2 Coronal images of three clinical examples. *Top row*: two PET/CT studies showing various degrees of uptake heterogeneity in large lung lesions. *Middle row*: PET response study of a subject with lung cancer. *Bottom row*: PET/CT images of a patient with metastatic liver lesions before and after treatment
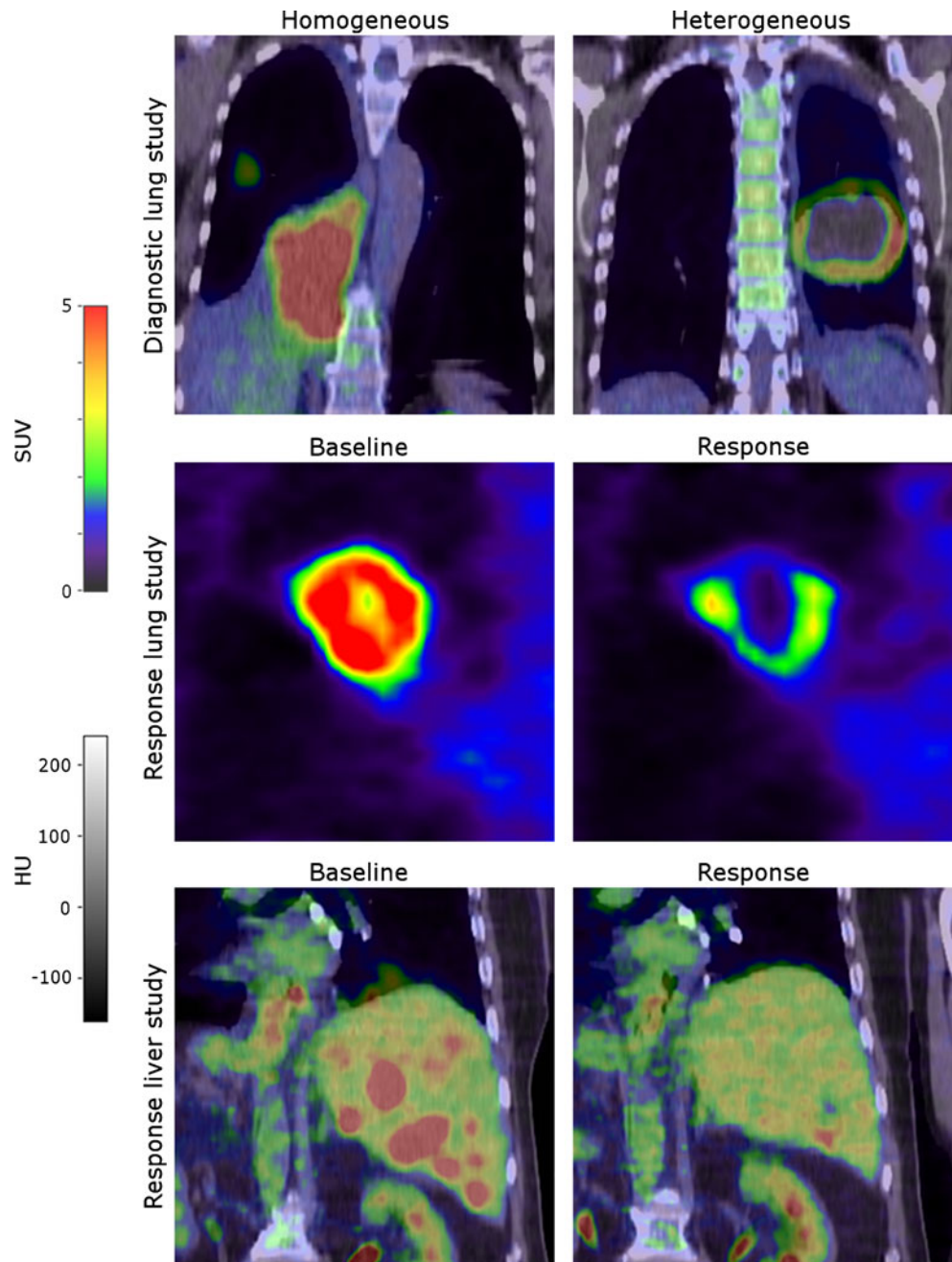
images were then forward-projected. Partial volume effects were introduced by smoothing the obtained sinograms using a 5-mm full-width at half-maximum (FWHM) Gaussian kernel. Poisson noise was added to the emission sinograms obtained. These sinograms were then reconstructed using normalization and attenuation weighted ordered subsets expectation maximization (NAW-OSEM) with 4 iterations and 16 subsets, and post-smoothed using a 5-mm FWHM Gaussian kernel. The resulting noise level (COV of voxel values within a VOI ~15%) was similar to that observed in PET studies acquired on a PET/CT scanner (Gemini TF 64, Philips Healthcare, Cleveland, OH, USA) [23]. All reconstructed images consisted of 30 planes of 256×256 voxels with a voxel size of 2.56×2.56×2.56 mm³. For each tumour type, 100 noisy simulation images were generated to investigate accuracy and precision of the various parameters as described above.

Human studies

Four different human FDG scans were included (Fig. 2). Two subjects were acquired on a whole-body PET/CT scanner (Gemini TF 64, Philips Healthcare, Cleveland, OH, USA). One subject (male, 78 kg, 59 years, 188 MBq FDG injection) had a primary lung tumour with heterogeneous uptake, whilst the other (female, 70 kg, 75 years, 184 MBq FDG injection) had a primary lung tumour with relative homogeneous uptake. In addition, one subject with a primary lung tumour (female, 60 years; baseline scan: 57 kg, 163 MBq FDG injection; response scan: 54 kg, 163 MBq FDG injection) was acquired twice (one scan before and one scan after one course of chemotherapy) on a whole-body PET scanner (ECAT EXACT HR+, CTI/Siemens, Knoxville, TN, USA). Furthermore, one subject with advanced liver metastases of a gastrointestinal malignancy (male, 65 years; baseline scan: 71 kg, 566 MBq FDG injection; response scan: 73 kg, 551 MBq FDG injection) was acquired twice (one scan before and one scan after three courses of chemotherapy) on a whole-body PET/CT scanner (Biograph, CTI/Siemens, Knoxville, TN, USA). All PET and CT data were collected as part of ongoing clinical studies, which were approved by an authorized medical Ethics Review Committee, and informed consent was obtained from each patient prior to inclusion in the study.

Data analysis

For all simulations and human scans, mean±SD were calculated for all parameters under investigation. In the case of the simulations, VOIs were taken from baseline (VOI^{BL}) scans or redrawn on response (VOI^R) scans. For the human scans, VOIs were drawn using a 50% threshold isocontour method on baseline PET(/CT) scans and these were then transformed manually to response scans to obtain a repositioned VOI^{BL}. In addition, VOIs were drawn using a 50% threshold isocontour method on the response scans (VOI^R). All uptake (heterogeneity) parameters were calculated for all VOIs.

Results

Simulations

Noise reduction and partial volume correction

AUC-CSH data obtained from original (i.e. noise and partial volume free), simulated PET (i.e. with noise and partial volume effects) and simulated PET with additional use of bilateral filter and Van Cittert deconvolution (PET+ BF+CIT) are shown in Table 2. A selection of responses is shown in Fig. 3. Table 3 illustrates that PET+BF+CIT gave
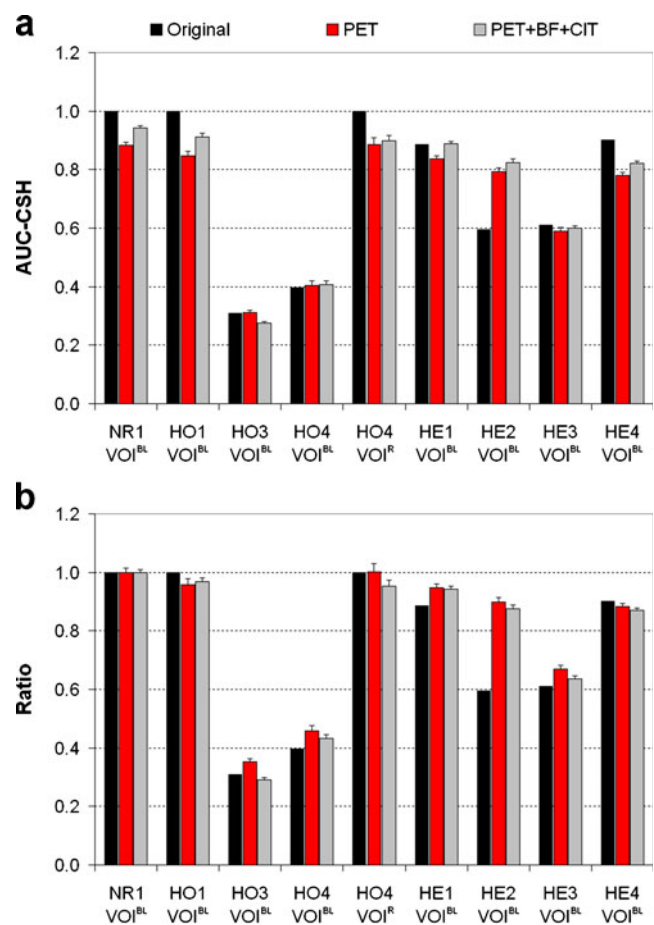

Fig. 3 AUC-CSH (**a**) and ratio (**b**) for various types of responses derived from simulated original (noise and partial volume free), PET and PET+BF+CIT scans for simulation 1. The ratio was obtained by dividing AUC-CSH of the response scan by that of the baseline scan

Table 3 AUC-CSH values of different tumour responses for simulated original (i.e. noise and partial volume free), PET and PET+BF+CIT[a] images

Response tumour type	Code[b]	VOI[c]	AUC-CSH		
			Original	PET	PET+BF+CIT
Simulation 1					
No response	NR1	VOI[BL]	1.00	0.88	0.94
Homogeneous	HO1	VOI[BL]	1.00	0.85	0.91
	HO2	VOI[BL]	1.00	0.91	0.96
	HO3	VOI[BL]	0.31	0.31	0.27
		(VOI[R])	(1.00)	(0.89)	(0.89)
	HO4	VOI[BL]	0.40	0.40	0.41
		(VOI[R])	(1.00)	(0.89)	(0.90)
	HO5	VOI[BL]	0.27	0.27	0.22
		(VOI[R])	(1.00)	(0.89)	(0.88)
Heterogeneous	HE1	VOI[BL]	0.89	0.84	0.89
	HE2	VOI[BL]	0.60	0.79	0.82
	HE3	VOI[BL]	0.61	0.59	0.60
	HE4	VOI[BL]	0.90	0.78	0.82
Simulation 2					
No response	NR2	VOI[BL]	0.61	0.59	0.60
Homogeneous	HOA	VOI[BL]	0.27	0.27	0.22
		(VOI[R])	(1.00)	(0.89)	(0.88)
	HOB	VOI[BL]	1.00	0.91	0.96
Heterogeneous	HEA	VOI[BL]	0.89	0.84	0.89
	HEB	VOI[BL]	0.61	0.59	0.60

[a] Simulated PET with additional use of bilateral filter and Van Cittert deconvolution

[b] Description of the various simulated tumour responses are given in Tables 1 (simulation 1) and 2 (simulation 2)

[c] VOI either defined on baseline (VOI[BL]) or response (VOI[R]) scans. VOI[R] is only provided when the volume of VOI[R] is different from the volume of VOI[BL]

a significant improvement in AUC-CSH compared to PET, at least when AUC-CSH was normalized to the baseline AUC-CSH (Student's t test, $p < 0.05$). Nevertheless, the actual improvement was rather modest and not significant for non-normalized data ($p = 0.10$). Figure 4 shows that CSH curves improve visually after partial volume correction and noise reduction. More importantly, after image denoising and partial volume correction, AUC-CSH varied with the degree of heterogeneity, i.e. lower AUC-CSH corresponds with a visually more heterogeneous tracer distribution as well as with the variability of voxel values

within a VOI. Therefore, in the remainder of this article, only CSH data for PET+BF+CIT images will be provided.

Simulating various response types

Table 4 shows ratios of AUC-CSH (obtained from PET+BF+CIT images) and several other parameters (obtained from PET) for various response types. A selection of responses is shown in Figs. 5 and 6 for simulations 1 and 2, respectively. These figures also show the effects of various SUV and volume normalizations on AUC-CSH. Both
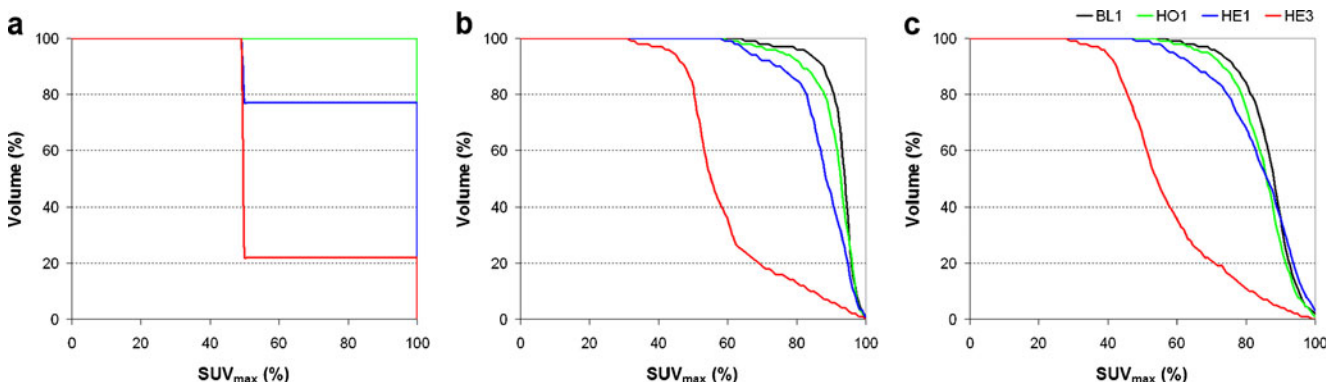

Fig. 4 CSHs for baseline and typical homogeneous and heterogeneous responses obtained derived from simulated original (**a**), PET (**b**) and PET+BF+CIT (**c**) scans for simulation 1. Note that HO1 overlaps with BL1 in (**a**)

Table 4 Ratios of AUC-CSH (derived from PET+BF+CIT[a]), SUV$_{max}$, SUV$_{mean}$ 1/SD and 1/COV (derived from PET) for different tumour responses. All values are normalized to the baseline scan

Response tumour type	Code[b]	VOI[c]	AUC-CSH	AUC-CSH^S	AUC-CSH^V	AUC-CSH^{SV}	SUV$_{max}$	SUV$_{mean}$	1/SD	1/COV	Volume
Simulation 1											
No response	NR1	VOI^{BL}	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Homogeneous	HO1	VOI^{BL}	0.97	0.49	0.97	0.49	0.52	0.50	3.13	0.88	1.00
	HO2	VOI^{BL}	1.02	2.04	1.02	2.04	1.96	2.01	0.31	1.12	1.00
	HO3	VOI^{BL}	0.29	0.26	0.29	0.26	0.89	0.31	0.14	0.12	1.00
		(VOI^R)	(0.94)	(0.83)	(0.12)	(0.10)	(0.89)	(0.90)	(0.64)	(0.72)	(0.13)
	HO4	VOI^{BL}	0.43	0.18	0.43	0.18	0.44	0.20	0.81	0.18	1.00
		(VOI^R)	(0.95)	(0.39)	(0.12)	(0.05)	(0.44)	(0.44)	(2.80)	(0.74)	(0.13)
	HO5	VOI^{BL}	0.23	0.43	0.23	0.43	1.81	0.55	0.03	0.09	1.00
		(VOI^R)	(0.93)	(1.72)	(0.12)	(0.22)	(1.81)	(1.84)	(0.16)	(0.74)	(0.13)
Heterogeneous	HE1	VOI^{BL}	0.94	0.89	0.94	0.89	0.93	0.88	0.50	0.63	1.00
	HE2	VOI^{BL}	0.88	0.65	0.88	0.65	0.73	0.65	0.88	0.61	1.00
	HE3	VOI^{BL}	0.64	1.18	0.64	1.18	1.84	1.23	0.07	0.34	1.00
	HE4	VOI^{BL}	0.87	1.71	0.87	1.71	1.93	1.70	0.07	0.44	1.00
Simulation 2											
No response	NR2	VOI^{BL}	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Homogeneous	HOA	VOI^{BL}	0.37	0.36	0.37	0.36	0.98	0.44	0.40	0.28	1.00
		(VOI^R)	(1.47)	(1.46)	(0.18)	(0.18)	(0.98)	(1.49)	(2.15)	(2.19)	(0.13)
	HOB	VOI^{BL}	1.60	1.73	1.60	1.73	1.06	1.63	4.13	3.32	1.00
Heterogeneous	HEA	VOI^{BL}	1.48	0.75	1.48	0.75	0.51	0.72	6.72	1.86	1.00
	HEB	VOI^{BL}	1.01	0.49	1.01	0.49	0.50	0.50	4.05	1.00	1.00

[a] Simulated PET with additional use of bilateral filter and Van Cittert deconvolution

[b] Description of the various simulated tumour responses are given in Tables 1 (simulation 1) and 2 (simulation 2)

[c] VOI either defined on baseline (VOI^{BL}) or response (VOI^R) scans. VOI^R is only provided when the volume of VOI^R is different from the volume of VOI^{BL}

Table 4 and Figs. 5 and 6 show that normalized values of AUC-CSH and 1/COV corresponded well with the level of response heterogeneity, which was not seen using SUV$_{max}$, SUV$_{mean}$ or 1/SD alone.

Simulation 1 (Fig. 5) shows that for homogeneous responses, normalized (ratio to the baseline) values of AUC-CSH and 1/COV ranged from 0.93 to 1.02 and from 0.72 to 1.12, respectively, while non-normalized values ranged from 0.82 to 0.96 and from 0.08 to 0.13, respectively. For heterogeneous responses, normalized values of AUC-CSH and 1/COV ranged from 0.64 to 0.94 and from 0.34 to 0.63, respectively, whilst non-normalized values ranged from 0.60 to 0.89 and from 0.04 to 0.07, respectively. AUC-CSH^S showed an increase or decrease when SUV$_{max}$ increased or decreased, respectively. Similarly, AUC-CSH^V showed an increase or decrease when volume increased or decreased, respectively. AUC-CSH^{SV} showed the combined effect of a change in volume and a change in SUV$_{max}$. 1/COV did not correspond with tumour heterogeneity when data were obtained from PET+BF+CIT images, where ranges were between 0.45 and 1.11 and between 0.26 and 0.62 for homogeneous and heterogeneous responses, respectively.

Simulation 2 (Fig. 6) also shows that AUC-CSH and 1/COV of a homogeneous response (HEB, 1.01 and 1.00, respectively), normalized to the baseline, falls within the same range as for simulation 1. Responses that show an increase in homogeneity showed an increase in normalized AUC-CSH and 1/COV, ranging from 1.47 to 1.60 and from 1.86 to 3.32, respectively. For non-normalized AUC-CSH and 1/COV, this trend was not observed (increase in homogeneity: 0.88–0.96 and 0.07–0.13, respectively; homogeneous response: 0.60 and 0.04, respectively).

Human studies

Figure 7 shows CSHs for the four clinical scans shown in Fig. 2. Values of the various parameters are shown in Table 5. For the diagnostic lung studies, AUC-CSH correctly indicated heterogeneity for the tumour with more heterogeneous uptake (0.47) and a more homogeneous distribution for the homogeneous one (0.79). For the response studies, AUC-CSH indicated an increase in heterogeneity (AUC-CSH: 0.74 to 0.63) for the lung tumour (Fig. 2, second row), whilst a decrease in
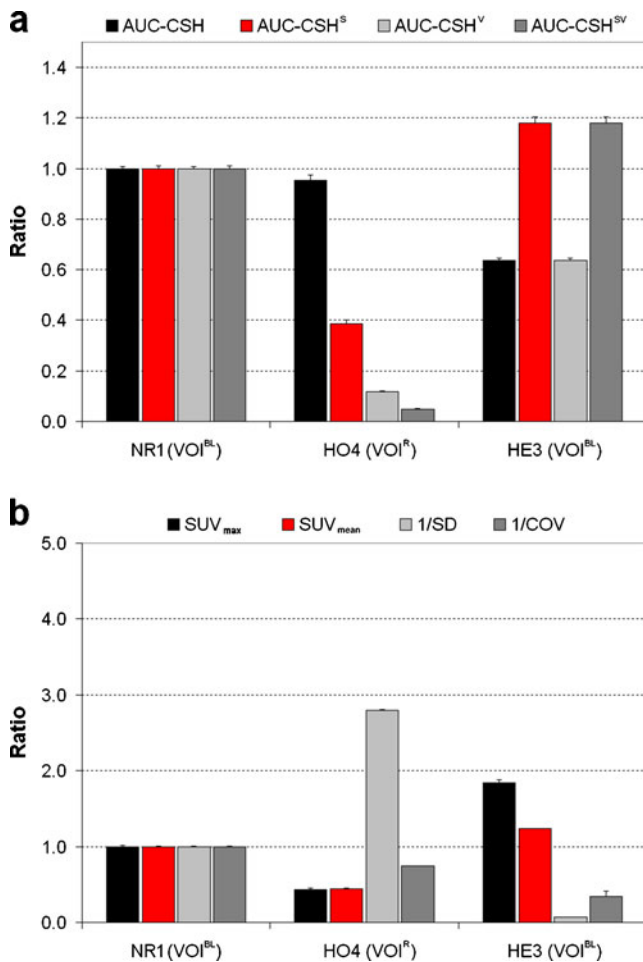
Fig. 5 Ratio of various types of AUC-CSH (**a**) and SUV$_{max}$, SUV$_{mean}$, 1/SD and 1/COV (**b**) for various types of responses for simulation 1. AUC-CSH and other uptake parameters were derived from simulated PET+BF+CIT and PET response images, respectively. The ratio for each parameter was obtained by dividing its value from the response scan by that from the baseline scan

tissue [8]. Although partial volume effects are unavoidable due to the limited resolution of PET, they need to be corrected for as much as possible. Recently, Hoetjes et al. [20] showed that image-based Van Cittert deconvolution is useful for partial volume correction in oncology. The results of the present study showed that a Van Cittert deconvolution with Gibbs prior, in combination with bilateral filtering to limit spatial noise, did improve CSH curves visually (Fig. 4) and improved accuracy of AUC-CSH (Fig. 3 and Table 3).

The simulations showed that both AUC-CSH and 1/COV corresponded well with level of tumour heterogeneity or heterogeneous response, in contrast to SUV$_{max}$, SUV$_{mean}$ or 1/SD (Fig. 5 and Table 4). This means that a change in AUC-CSH corresponded with a visual apparent change in tracer uptake distribution as well as with the variability of voxel values within a tumour. Although the latter can be expressed by COV (%) of voxel values, AUC-CSH might
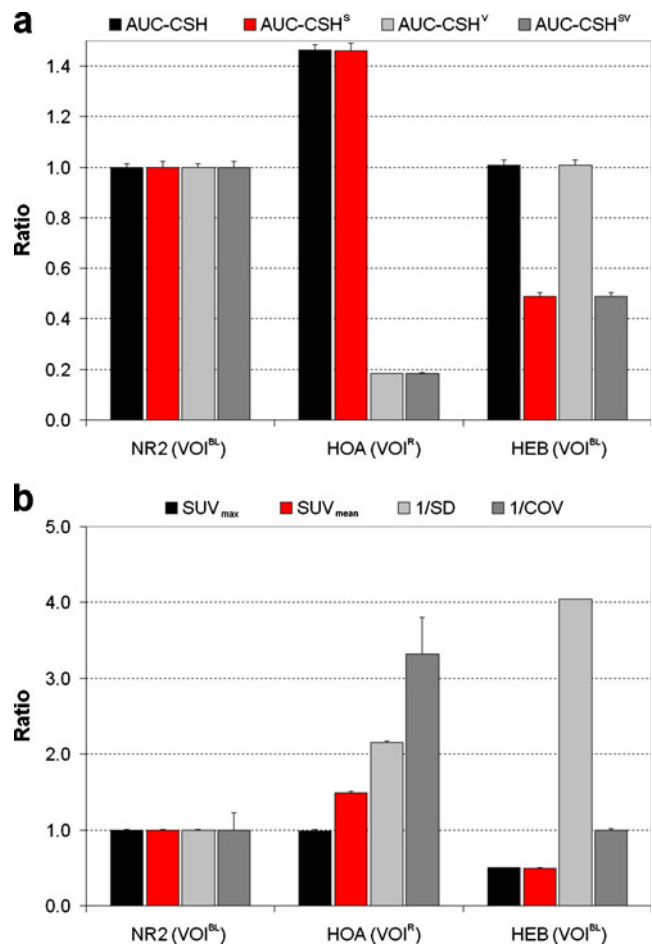

Fig. 6 Ratio of various types of AUC-CSH (**a**) and SUV$_{max}$, SUV$_{mean}$, 1/SD and 1/COV (**b**) for various types of responses for simulation 2. AUC-CSH and other uptake parameters were derived from simulated PET+BF+CIT and PET response images, respectively. The ratio for each parameter was obtained by dividing its value from the response scan by that from the baseline scan

heterogeneity (AUC-CSH: 0.18 to 0.40) was seen for the liver metastases (Fig. 2, second row). All the increases and decreases of AUC-CSH were in agreement with visual interpretation. Similar trends were observed for 1/COV.

Discussion

Both partial volume effects and image noise will lead to apparent tumour heterogeneity, and a correction for these confounding effects is needed for accurate calculation of measures that quantify tumour heterogeneity. Partial volume effects could prevent identification of the origin of counts in voxels, as these consist of average values of which the origin is not always known. For example, low uptake of FDG in a voxel could be the result of well oxygenated tumour tissue (with low and uniform FDG uptake) or averaging over a region with both hypoxic (high FDG uptake) and necrotic (low FDG uptake)
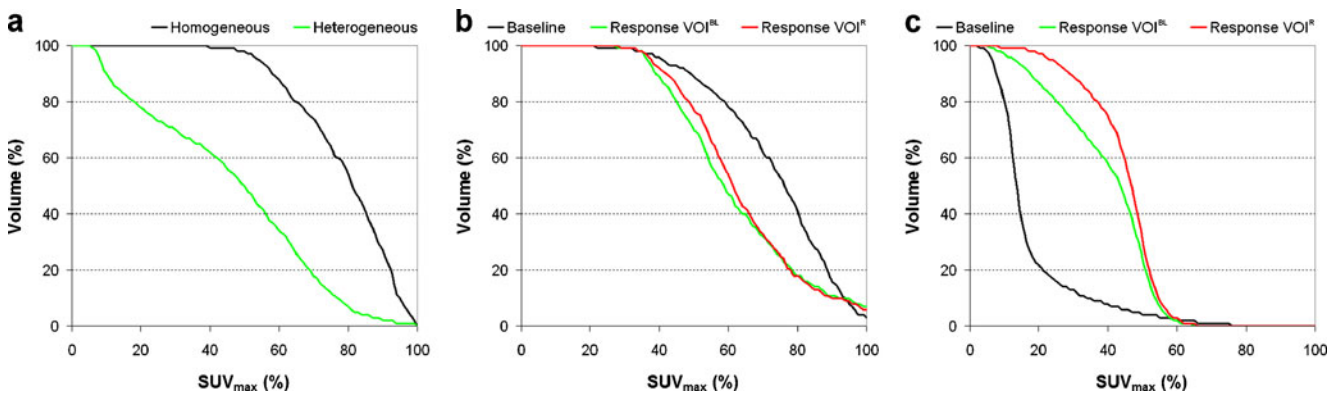
Fig. 7 CSHs for diagnostic lung study (**a**), response lung study (**b**) and response liver study (**c**). VOIs of response scans were either defined on the baseline scan (VOI^{BL}) or on the response scan (VOI^R).

The AUC-CSH and/or its change corresponded well with various degrees of tracer uptake heterogeneity and/or its change

have the additional advantage of exploring changes in tracer uptake distribution relative to baseline distribution, i.e. to normalize the x- and y-axis to baseline volume and SUV$_{max}$, respectively. In practice, this means that 100% corresponds to the metabolic volume (y-axis) and SUV$_{max}$ (x-axis) observed in the baseline study. When such normalization is applied, tumour growth would result in a CSH that goes beyond the 100% on the y-axis. Likewise, an increase of SUV$_{max}$ means that the CSH would go beyond the 100% of the x-axis. Thus, visual representation of CSH curves for baseline and response studies in one plot can then assist in a quick interpretation of both metabolic volume and tracer uptake changes. Note that a change in AUC-CSH, when both axes are normalized to the baseline, equals that of the total lesion glycolysis (TLG, calculated as the product of SUV$_{mean}$ and volume). However, a constant TLG could still be seen when tumour size has increased with a corresponding decrease in SUV. Use of normalized CSH would also demonstrate a constant AUC-CSH (Fig. 8), but would provide CSH curves with different shapes. By not renormalizing CSH (i.e. CSH is generated using the tumour volume

and SUV$_{max}$ of the image being analysed) an index for tracer heterogeneity or tracer uptake variability over the tumour at that time is obtained. These different normalizations may be useful for response monitoring, as they provide additional information on the response. The type of normalization of AUC-CSH that would be clinically most relevant needs to be further assessed in future studies.

CSH could be used for various applications. The focus of the present study was measurement of heterogeneity in FDG uptake in an NSCLC tumour. Two clinical examples (Figs. 2 and 7 and Table 5) indicate that AUC-CSH can quantify the degree of heterogeneity in FDG uptake, both at baseline (diagnostic studies) and as a result of therapy (response studies). For response studies, VOIs could be defined on either baseline or response scans. Definition of VOI on the baseline and response scans might identify increase in necrotic tissue tumour after treatment. Definition of VOI on the response scan would indicate whether overall heterogeneity in FDG uptake of the tumour had changed in the presence of a volume change. The definition that would be clinically most relevant needs to be further assessed in future

Table 5 AUC-CSH, SUV$_{max}$, SUV$_{mean}$ 1/SD and 1/COV data for the four human scans shown in Fig. 2, as obtained from PET+BF+CIT[a] images

Study	Type	VOI[b]	AUC-CSH	AUC-CSH^S	AUC-CSH^V	AUC-CSH^{SV}	SUV$_{max}$	SUV$_{mean}$	1/SD	1/COV	Volume (ml)
Diagnostic lung study	Homogeneous	VOI^{BL}	0.79	6.7	93	780	13.2	10.5	0.48	0.06	120
	Heterogeneous	VOI^{BL}	0.47	2.0	91	390	5.9	2.8	0.64	0.02	200
Response lung study	Baseline	VOI^{BL}	0.74	6.7	28	250	9.1	6.7	0.41	0.043	37
	Response	VOI^{BL}	0.63	1.9	23	71	3.1	1.9	2.6	0.031	37
		(VOI^R)	(0.64)	(2.0)	(28)	(84)	(3.1)	(2.0)	(3.0)	(0.034)	(43)
Response liver study	Baseline	VOI^{BL}	0.18	3.9	370	7,900	38	6.7	0.064	0.013	2,000
	Response	VOI^{BL}	0.40	2.3	800	4,600	7.9	3.2	1.2	0.028	2,000
		(VOI^R)	(0.45)	(2.6)	(710)	(4,100)	(7.9)	(3.6)	(2.1)	(0.044)	(1,600)

[a] Simulated PET with additional use of bilateral filter and Van Cittert deconvolution

[b] VOI either defined on baseline (VOI^{BL}) or response (VOI^R) scans. VOI^R is only provided for response studies
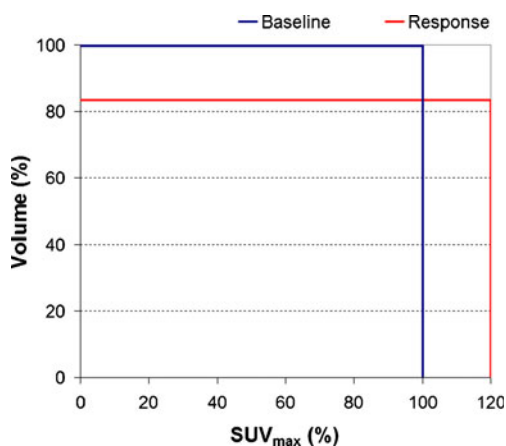
Fig. 8 CSHs for theoretical baseline and response study. Both show the same AUC, but different curves

studies. Another application might be assessment of response across multiple metastatic lesions, such as in the liver. The liver often contains many metastatic lesions, which are difficult to monitor individually. In these cases AUC-CSH could be used to characterize a global change in tumour load in combination with e.g. TLG [24]. Finally, CSH and AUC-CSH may be used to characterize differences between CT-based anatomical and PET-based metabolic tumour volumes and/or tracer distributions, i.e. a high AUC-CSH would indicate a good correspondence between CT tumour volume and tracer distribution, whilst a low AUC-CSH indicates a large discrepancy between CT tumour volume and PET tracer distribution.

To the best of our knowledge, this is the first study that attempts to characterize and quantify heterogeneity of tracer uptake using the AUC-CSH and different normalizations. The performance of the proposed method was assessed using simulations, and a few clinical cases were used to illustrate the performance of the method and different normalizations. Future studies need to address the potential clinical value of CSH and are aimed at test-retest studies as well as the application of the method on larger clinical data sets. Furthermore, it would be ideal to have a database or benchmark data set containing more realistically simulated anthropomorphic PET images based on e.g. Monte Carlo simulations [25, 26]. This would allow one to not only more extensively explore uptake heterogeneity measures but also to validate new tumour segmentation algorithms. The proposed method for parameterizing uptake heterogeneity is not spatially invariant, unlike the method proposed by O'Sullivan et al. [12]. For patients with sarcoma, this method that assesses spatial intratumoural heterogeneity in FDG uptake has been shown to be a predictor of patient outcome [27]. However, preliminary studies already indicated that other measures derived from the CSH, such as the % of volume derived at a fixed % of SUV or % of SUV derived at

a fixed % of volume, could be used as a prognostic factor in NSCLC [28] and head and neck cancer [14]. In addition, further studies are needed to assess the value of the different methods to normalize these CSH to baseline values (SUV and/or volume). Nevertheless, the present preliminary results might indicate the potential value of CSH and AUC-CSH characterizing metabolic tumour heterogeneity.

Conclusion

These initial results indicate that AUC-CSH might be used as a quantitative index of heterogeneity in tracer uptake and it can be used as a means to address (changes in) heterogeneity in response assessment studies. In addition, the results show that a Van Cittert deconvolution with Gibbs prior, in combination with bilateral filtering to limit spatial noise, improves the CSH curves visually.

References

1. Stroobants SG, D'Hoore I, Dooms C, De Leyn PR, Dupont PJ, De Wever W, et al. Additional value of whole-body fluorodeoxyglucose positron emission tomography in the detection of distant metastases of non-small-cell lung cancer. Clin Lung Cancer 2003;4:242–7.
2. Nestle U, Kremp S, Grosu AL. Practical integration of [18F]-FDG-PET and PET-CT in the planning of radiotherapy for non-small cell lung cancer (NSCLC): the technical basis, ICRU-target volumes, problems, perspectives. Radiother Oncol 2006;81:209–25.
3. Sørensen M, Horsman MR, Cumming P, Munk OL, Keiding S. Effect of intratumoral heterogeneity in oxygenation status on FMISO PET, autoradiography, and electrode Po2 measurements in murine tumors. Int J Radiat Oncol Biol Phys 2005;62:854–61.
4. Avril N, Menzel M, Dose J, Schelling M, Weber W, Jänicke F, et al. Glucose metabolism of breast cancer assessed by 18F-FDG PET: histologic and immunohistochemical tissue analysis. J Nucl Med 2001;42:9–16.
5. Zasadny KR, Tatsumi M, Wahl RL. FDG metabolism and uptake versus blood flow in women with untreated primary breast cancers. Eur J Nucl Med Mol Imaging 2003;30:274–80.
6. Tateishi U, Nishihara H, Tsukamoto E, Morikawa T, Tamaki N, Miyasaka K. Lung tumors evaluated with FDG-PET and dynamic CT: the relationship between vascular density and glucose metabolism. J Comput Assist Tomogr 2002;26:185–90.

7. Zhao S, Kuge Y, Mochizuki T, Takahashi T, Nakada K, Sato M, et al. Biologic correlates of intratumoral heterogeneity in 18F-FDG distribution with regional expression of glucose transporters and hexokinase-II in experimental tumor. J Nucl Med 2005;46:675–82.

8. Pugachev A, Ruan S, Carlin S, Larson SM, Campa J, Ling CC, et al. Dependence of FDG uptake on tumor microenvironment. Int J Radiat Oncol Biol Phys 2005;62:545–53.

9. van Baardwijk A, Bosmans G, van Suylen RJ, van Kroonenburgh M, Hochstenbag M, Geskes G, et al. Correlation of intra-tumour heterogeneity on 18F-FDG PET with pathologic features in non-small cell lung cancer: a feasibility study. Radiother Oncol 2008;87:55–8.

10. 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;45:1519–27.

11. van Baardwijk A, Dooms C, van Suylen RJ, Verbeken E, Hochstenbag M, Dehing-Oberije C, et al. The maximum uptake of (18)F-deoxyglucose on positron emission tomography scan correlates with survival, hypoxia inducible factor-1alpha and GLUT-1 in non-small cell lung cancer. Eur J Cancer 2007;43:1392–8.

12. O'Sullivan F, Roy S, Eary J. A statistical measure of tissue heterogeneity with application to 3D PET sarcoma data. Biostatistics 2003;4:433–48.

13. O'Sullivan F, Roy S, O'Sullivan J, Vernon C, Eary J. Incorporation of tumor shape into an assessment of spatial heterogeneity for human sarcomas imaged with FDG-PET. Biostatistics 2005;6:293–301.

14. El Naqa I, Grigsby P, Apte A, Kidd E, Donnelly E, Khullar D, et al. Exploring feature-based approaches in PET images for predicting cancer treatment outcomes. Pattern Recognit 2009;42:1162–71.

15. Drzymala RE, Mohan R, Brewster L, Chu J, Goitein M, Harms W, et al. Dose-volume histograms. Int J Radiat Oncol Biol Phys 1991;21:71–8.

16. Nestle U, Kremp S, Schaefer-Schuler A, Sebastian-Welsch C, Hellwig D, Rübe C, et al. Comparison of different methods for delineation of 18F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer. J Nucl Med 2005;46:1342–8.

17. van Velden FHP, Cheebsumon P, Yaqub M, Hoekstra OS, Lammertsma AA, Boellaard R. Evaluation of a cumulative SUV-volume histogram method for parameterizing heterogeneous tumour FDG uptake in PET studies. Eur J Nucl Med Mol Imaging 2010;37:S262.

18. Boellaard R, Cheebsumon P, van Velden FHP, Yaqub M, Lammertsma AA, Hoekstra OS. A novel cumulative SUV-volume histogram method for parameterizing heterogeneous tumour tracer uptake in oncology FDG PET studies. Eur J Nucl Med Mol Imaging 2010;37:S261.

19. Teo BK, Seo Y, Bacharach SL, Carrasquillo JA, Libutti SK, Shukla H, et al. Partial-volume correction in PET: validation of an iterative postreconstruction method with phantom and patient data. J Nucl Med 2007;48:802–10.

20. Hoetjes NJ, van Velden FHP, Hoekstra OS, Hoekstra CJ, Krak NC, Lammertsma AA, et al. Partial volume correction strategies for quantitative FDG PET in oncology. Eur J Nucl Med Mol Imaging 2010;37:1679–87.

21. Tohka J, Reilhac A. A Monte Carlo study of deconvolution algorithms for partial volume correction in quantitative PET. IEEE Nucl Sci Symp Conf Rec 2006;3339–3345.

22. Geets X, Lee JA, Bol A, Lonneux M, Grégoire V. A gradient-based method for segmenting FDG-PET images: methodology and validation. Eur J Nucl Med Mol Imaging 2007;34:1427–38.

23. Surti S, Kuhn A, Werner ME, Perkins AE, Kolthammer J, Karp JS. Performance of Philips Gemini TF PET/CT scanner with special consideration for its time-of-flight imaging capabilities. J Nucl Med 2007;48:471–80.

24. Larson SM, 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.

25. Jan S, Benoit D, Becheva E, Carlier T, Cassol F, Descourt P, et al. GATE V6: a major enhancement of the GATE simulation platform enabling modelling of CT and radiotherapy. Phys Med Biol 2011;56:881–901.

26. Segars WP, Sturgeon G, Mendonca S, Grimes J, Tsui BM. 4D XCAT phantom for multimodality imaging research. Med Phys 2010;37:4902–15.

27. Eary JF, O'Sullivan F, O'Sullivan J, Conrad EU. Spatial heterogeneity in sarcoma 18F-FDG uptake as a predictor of patient outcome. J Nucl Med 2008;49:1973–9.

28. El Naqa I, Bradley J, Guild C, Apte A, Fergus S, Dehdashti F, et al. Pattern recognition analysis of FDG-PET uptake characteristics for assessing response in NSCLC post-radiotherapy treatment. Int J Radiat Oncol Biol Phys 2008;72:S113.