ORIGINAL RESEARCH

Open Access

Volume of interest delineation techniques for ^{18}F-FDG PET-CT scans during neoadjuvant extremity soft tissue sarcoma treatment in adults: a feasibility study

Marc G. Stevenson[1], Lukas B. Been[1], Harald J. Hoekstra[1], Albert J. H. Suurmeijer[2], Ronald Boellaard[3,4] and Adrienne H. Brouwers[3*]

Abstract

Background: This study explores various volume of interest (VOI) delineation techniques for fluorine-18-fluorodeoxyglucose positron emission tomography with computed tomography (^{18}F-FDG PET-CT) scans during neoadjuvant extremity soft tissue sarcoma (ESTS) treatment.

Results: During neoadjuvant treatment, hyperthermic isolated limb perfusion (HILP) and preoperative external beam radiotherapy (EBRT), 11 patients underwent three ^{18}F-FDG PET-CT scans. The first scan was made prior to the HILP, the second after the HILP but prior to the start of the EBRT, and the third prior to surgical resection. An automatically drawn VOI$_{auto}$, a manually drawn VOI$_{man}$, and two gradient-based semi-automatically drawn VOIs (VOI$_{grad}$ and VOI$_{grad+}$) were obtained. Maximum standardized uptake value (SUV$_{max}$), SUV$_{peak}$, SUV$_{mean}$, metabolically active tumor volume (MATV), and total lesion glycolysis (TLG) were calculated from each VOI. The correlation and level of agreement between VOI delineation techniques was explored. Lastly, the changes in metabolic tumor activity were related to the histopathologic response. The strongest correlation and an acceptable level of agreement was found between the VOI$_{man}$ and the VOI$_{grad+}$ delineation techniques. A decline (VOI$_{man}$) in SUVmax, SUVpeak, SUVmean, TLG, and MATV (all $p < 0.05$) was found between the three scans. A > 75% decline in TLG between scan 1 and scan 3 possibly identifies histopathologic response.

Conclusions: The VOI$_{grad+}$ delineation technique was identified as most reliable considering reproducibility when compared with the other VOI delineation techniques during the multimodality neoadjuvant treatment of locally advanced ESTS. A significant decline in metabolic tumor activity during the treatment was found. TLG deserves further exploration as predictor for histopathologic response after multimodality ESTS treatment.

Keywords: Soft tissue sarcoma, ^{18}F-FDG PET-CT, Limb perfusion, Preoperative radiotherapy

Background

Soft tissue sarcomas (STS) are relatively rare malignancies, accounting for less than 1% of all cancers in adults. The number of patients presenting with STS each year is 600–700 in the Netherlands, leading to approximately 300 STS related deaths annually [1, 2].

Roughly 50–60% of the STS arise in the extremities [3, 4]. At presentation, some of these extremity soft tissue sarcomas (ESTS) are considered non-resectable or "locally advanced." Since the 1990s, neoadjuvant hyperthermic isolated limb perfusion (HILP) has been used in Europe to prevent limb amputation in these patients [5], resulting in a limb salvage rate of 80–90% in locally advanced ESTS nowadays [6–9]. HILP is used in all types of adult locally advanced ESTS. It allows to administer regional chemotherapy in high doses, as the affected limb is isolated from the systemic circulation

* Correspondence: a.h.brouwers@umcg.nl
[3]Department of Nuclear Medicine and Molecular Imaging, University Medical Center Groningen, University of Groningen, Hanzeplein 1, PO Box 30.001, 9700 RB Groningen, The Netherlands
Full list of author information is available at the end of the article

Stevenson et al. EJNMMI Research (2018) 8:42

Page 2 of 11

during the procedure. Neoadjuvant systemic chemotherapy in ESTS is currently under ongoing investigation, as the data available considering patients' oncological outcome are inconsistent [10–12].

Fluorine-18-fluorodeoxyglucose positron emission tomography with computed tomography (^{18}F-FDG PET-CT) scans have been used to evaluate tumor changes following HILP in locally advanced ESTS since the mid-1990s [13]. Pretreatment maximum standardized uptake value (SUV_{max}), metabolically active tumor volume (MATV), and total lesion glycolysis (TLG) were identified as significant predictors for overall survival in STS in a recent meta-analysis [14]. Furthermore, post-treatment SUV_{max} was shown to be promising in monitoring treatment response. However, the identification of this latter parameter was solely based on two articles included in this meta-analysis. The first only included rhabdomyosarcomas, which is a chemosensitive sarcoma, and the second only included chest wall sarcomas [14–16].

The SUV_{max} of a lesion depends solely on the highest measured ^{18}F-FDG uptake in one voxel, thereby making the measured SUV_{max} susceptible for noise [17]. Furthermore, the question remains whether this one measurement is representative for large, heterogeneous tumors, as STS. In contrast, the SUV_{max} is the most robust parameter when comparing various software delineation programs, delineation methods, and observers [18]. The outcome of MATV and TLG parameters are much more dependent of the method of tumor delineation and the software program used for these analyses. We hypothesized that the use of peak standardized uptake value (SUV_{peak}) and mean standardized uptake value (SUV_{mean}) in addition to SUV_{max}, TLG, and MATV might result in a more reliable prediction of tumor changes induced by neoadjuvant treatment.

To the best of our knowledge, the use of various VOI delineation techniques has not yet been explored in and during the neoadjuvant treatment of STS. Furthermore, in this patient population, no sequential analysis of multiple ^{18}F-FDG PET-CT scans has been performed previously. In this feasibility study, consecutive ^{18}F-FDG PET-CT scans per patient were used to investigate the use of four VOI delineation techniques because variations in VOI will directly affect the measured SUV_{mean}, MATV, and TLG and could thus affect the performance of the PET assessments. Furthermore, we explored the changes in metabolic tumor activity (SUV_{max}, SUV_{peak}, SUV_{mean}, MATV, and TLG) to neoadjuvant HILP and preoperative EBRT during the treatment course of locally advanced ESTS. Lastly, the relationship between changes in metabolic tumor activity and histopathologic response was explored.

Methods

This study has been approved by the Institutional Review Board (IRB), and the need for written informed consent was waived (IRB case number 2016.984). From 2011 to 2017, 11 patients with a median age of 64 (IQR 44–74; range 32–74) years were treated according to a novel treatment regimen consisting of neoadjuvant HILP, preoperative hypofractionated EBRT, followed by surgical resection of the tumor. All patients were diagnosed with a locally advanced, non-metastatic, high-grade ESTS (Table 1). Patients eligible for HILP treatment were included in this novel treatment regimen based on a tumor board decision. Inclusion and exclusion criteria, as well as treatment details, have been described in more detail elsewhere [19]. Patients were scheduled for three ^{18}F-FDG PET-CT scans. The first scan was made prior to the start of neoadjuvant treatment (baseline) and the second after the HILP, but prior to the start of the preoperative EBRT and was additionally used for EBRT delineation. The third scan was made after completion of the neoadjuvant treatment (HILP

Table 1 Patient and tumor characteristics

Patient number	Gender	Age (years)	Histopathologic findings	Tumor location	Tumor size (cm)
1	M	32	Synovial sarcoma	Upper leg	6
2	F	41	Synovial sarcoma	Lower leg	4
3	F	74	Pleomorphic undifferentiated sarcoma	Upper leg	10
4	M	54	Pleomorphic undifferentiated sarcoma	Upper leg	17
5	M	63	Pleomorphic undifferentiated sarcoma	Lower leg	9
6	M	71	Myxofibrosarcoma	Upper leg	5
7	M	44	Myxofibrosarcoma	Upper leg	17
8	M	74	Pleomorphic undifferentiated sarcoma	Knee	7
9	M	64	Leiomyosarcoma	Knee	6
10	M	75	Pleomorphic undifferentiated sarcoma	Lower leg	8
11	M	67	Leiomyosarcoma	Knee	6

Stevenson *et al. EJNMMI Research* (2018) 8:42

Page 3 of 11

and EBRT), but prior to surgical resection. Figure 1 illustrates the change in ^{18}F-FDG uptake during the treatment course for one of the patients.

^{18}F-FDG PET-CT

The ^{18}F-FDG PET-CT scans were performed using a hybrid PET-CT scanner (Siemens Biograph mCT). Patients fasted at least 6 h prior to scanning, and fasting glucose levels were checked at time of injection; none of the patients suffered from diabetes mellitus. ^{18}F-FDG (3 MBq/kg) was injected, and the PET-CT scan was started 1 h afterwards. Patients were scanned in supine position, and images of the affected limb were acquired in 3D mode, in two to five bed positions, 1–3 min/bed position based on the patient's body weight. A preceding low dose CT scan was performed and used for attenuation and scatter correction. All images were reconstructed using an EARL compliant protocol; from 2011 to 2014, the images were reconstructed using the following reconstruction: 3i_24s, image size 400, filter Gaussian, and FWHM 5.0 mm, and from 2014 to 2017, the images were reconstructed with the following reconstruction parameters: 3i_21s, image size 256, filter Gaussian, FWHM 6.5 mm, and quality ref. mAS 30. All scans were acquired according to European Association of Nuclear Medicine guidelines (version 1.0/2.0) [20, 21].
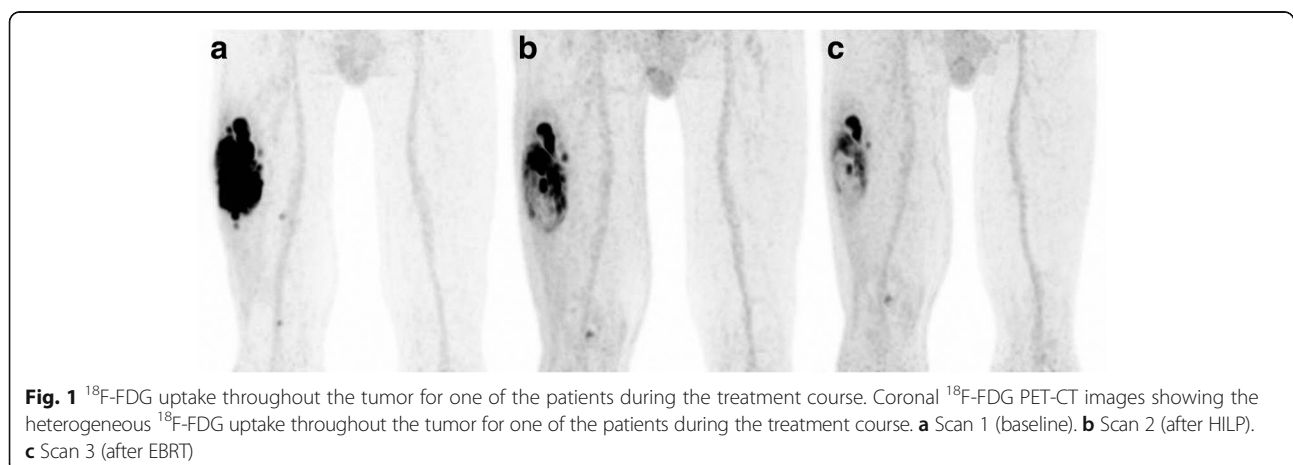
Image analyses

Scans were imported into Accurate (in-house developed analysis software, as previously used by Frings and Kramer et al. [22, 23]) and recently described by Boellaard [24]. Scans were reviewed and analyzed by one researcher. To explore the effect of various delineation techniques on the measurement of the metabolic parameters, the volume of interest (VOI) of each tumor was drawn in four different ways: (1) an automatically drawn VOI$_{auto}$ (using 50% of the SUV$_{peak}$ contour, corrected for local background [22]), (2) a manually drawn VOI$_{man}$ (visually following tumor contours), and (3) a semi-automatic drawn VOI$_{grad}$ (a contour that is located at the maximum PET image intensity gradient near the boundary of the tumor). Because of tumor heterogeneity, necrotic tumor parts (mostly tumor centers) were not included in this third VOI. Therefore, a fourth VOI was derived from the VOI$_{grad}$, in which all necrotic tumor parts were manually filled and included, resulting in the fourth VOI$_{grad+}$ (Fig. 2).

Five metabolic parameters, SUV$_{max}$ (voxel with the highest SUV value), SUV$_{peak}$ (using a 1 mL sphere), SUV$_{mean}$, TLG (SUV$_{mean}$ × MATV), and MATV, all based on lean body mass, as recommended by Boellaard et al. [21], were derived for the four VOI delineation techniques.

Due to tumor necrosis in most tumors, either treatment-induced or due to tumor heterogeneity, only the VOI$_{man}$ comprised the entire tumor (including necrosis). Therefore, the VOI$_{man}$ was chosen as reference measurement, and the other VOI techniques were compared with the VOI$_{man}$. We selected VOI$_{man}$ as reference VOI for pragmatic reasons (as the VOI$_{man}$ encompasses the entire tumor), not suggesting that this approach is best.

Correlation analyses, Bland-Altman analyses, and patient ranking were performed to compare correlation and level of agreement between the VOI delineation techniques. Bland-Altman analyses [25] and patient ranking are described in more detail in Additional file 1. Changes in metabolic tumor activity during neoadjuvant treatment were measured using the five metabolic parameters obtained from the reference VOI$_{man}$ and were related to histopathologic responses. Histopathologic tumor responses were established in accordance with the European Organization for Research and Treatment of Cancer-Soft Tissue and Bone Sarcoma Group (EORTC-STBSG) STS response score [19]. Grade A represents no stainable tumor cells, grade B single stainable
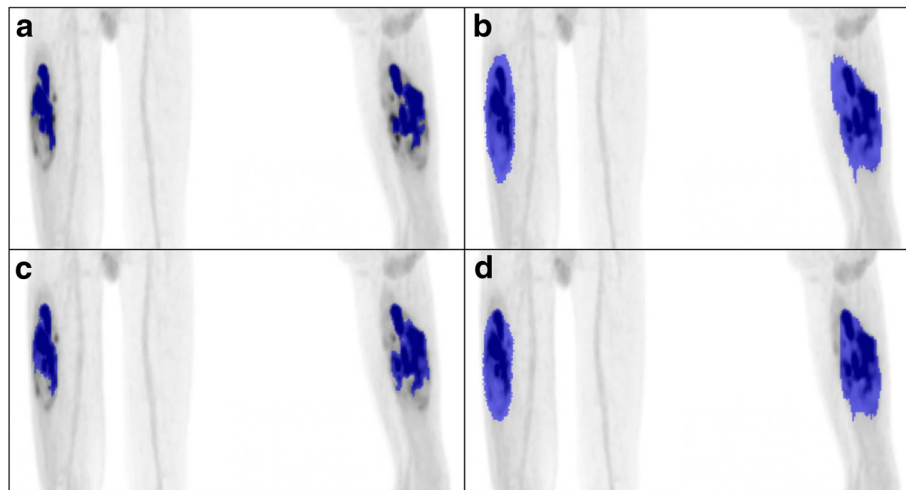

Fig. 1 ^{18}F-FDG uptake throughout the tumor for one of the patients during the treatment course. Coronal ^{18}F-FDG PET-CT images showing the heterogeneous ^{18}F-FDG uptake throughout the tumor for one of the patients during the treatment course. **a** Scan 1 (baseline). **b** Scan 2 (after HILP). **c** Scan 3 (after EBRT)

Stevenson *et al. EJNMMI Research* (2018) 8:42

Page 4 of 11

Fig. 2 Differences in tumor delineation between the four VOI delineation techniques. An example illustrating the differences in tumor delineation between the four VOI delineation techniques, for patient 4 scan 2. **a** VOI_{auto}. **b** VOI_{man}. **c** VOI_{grad}. **d** VOI_{grad+}

tumor cells or small clusters (overall below 1% of the whole specimen), grade C ≥ 1 to < 10% stainable tumor cells, grade D ≥ 10 to < 50% stainable tumor cells, and grade E ≥ 50% stainable tumor cells [26].

Histopathologic responders had tumor remnants which showed < 10% stainable cells, combining response grades A, B, and C. Non-responders had ≥ 10% stainable cells in their tumor remnant, grade D or E. Lastly, the relationship between changes in metabolic tumor activity and histopathologic responses was explored.

Statistical analysis

Discrete variables were summarized with frequencies and percentages and continuous variables with medians and interquartile ranges (IQRs); none of the variables were normally distributed. Fisher's exact and Mann-Whitney U test were used to compare variables. Wilcoxon signed rank and Friedman's test were used to compare the measurements between the three scans. Correlation coefficients were calculated and tested using Spearman's test. The level of agreement between VOI techniques was determined by Bland-Altman analyses [25]. A p value < 0.05 indicated statistical significance. Microsoft Excel (2010) was used to create the Bland-Altman plots. SPSS version 23.0 (IBM SPSS Statistics for Windows, Version 23.0 Armonk, NY: IBM Corp) and GraphPad Prism version 5.04 (GraphPad Software for Windows, San Diego California USA) were used for statistical analyses.

Results

Thirty-two ^{18}F-FDG PET-CT scans were acquired. The third PET-CT scan of patient 10 could not be performed

due to scheduling difficulties. For patient 1, in scan 3 it was not possible to draw a VOI_{auto}, since the tumor showed an almost complete metabolic response at this treatment stage and it did not meet the margin thresholds to complete the VOI_{auto}. Since it was possible to define the other three types of VOIs, this scan was included in the analyses and a value of zero was given to the metabolic parameters for the VOI_{auto}. The median time between the HILP and scan 2 was 21 [18–21] days, whereas the time between the end of EBRT and scan 3 was 3 (1–3) days.

Correlation, level of agreement, and ranking of patients between VOIs

The correlation between VOIs for all scans and all metabolic parameters was strongest between the VOI_{man} and the VOI_{grad+}, as indicated in gray in Table 2. The Bland-Altman plots showed an acceptable level of agreement between the VOI_{man} and the VOI_{grad+} (Additional file 2: Figure S1).

No larger difference than 1 place in ranking for SUV_{mean}, and TLG for the serial ^{18}F-FDG PET-CT scans was found when comparing the VOI_{man} and the VOI_{grad+} delineation techniques, for the MATV no larger difference than 2 places in ranking was found. A relative large difference of 4 or more in ranking between VOI delineation techniques is indicated in gray in Additional file 3: Table S1. Among others, this was found for the MATV at scan 1 of patient 7 with considerable necrotic tumor parts. The measured MATV was found to be highest when using the $VOI_{man, grad and grad+}$ techniques. However, when the VOI_{auto} technique was used, it was only ranked a 9th place due to exclusion of tumor necrosis.

Stevenson *et al. EJNMMI Research* (2018) 8:42

Page 5 of 11

Table 2 Spearman's correlation between the VOI_{man} and $VOI_{auto/grad/grad+}$ for the serial ^{18}F-FDG PET-CT scans

Parameter	Scan 1		Scan 2		Scan 3	
	Correlation coefficient	*p* value	Correlation coefficient	*p* value	Correlation coefficient	*p* value
SUV_{max}						
$VOI_{man-auto}$	1.000	NA	1.000	NA	0.988	< 0.001
$VOI_{man-grad}$	1.000	NA	1.000	NA	1.000	NA
$VOI_{man-grad+}$	1.000	NA	1.000	NA	1.000	NA
SUV_{peak}						
$VOI_{man-auto}$	1.000	NA	1.000	NA	0.988	< 0.001
$VOI_{man-grad}$	1.000	NA	1.000	NA	1.000	NA
$VOI_{man-grad+}$	1.000	NA	1.000	NA	1.000	NA
SUV_{mean}						
$VOI_{man-auto}$	0.964	< 0.001	0.836	0.001	0.564	0.090
$VOI_{man-grad}$	0.991	< 0.001	0.882	< 0.001	0.758	0.011
$VOI_{man-grad+}$	0.991	< 0.001	0.982	< 0.001	0.988	< 0.001
TLG						
$VOI_{man-auto}$	0.845	0.001	0.982	< 0.001	0.842	0.002
$VOI_{man-grad}$	0.991	< 0.001	1.000	NA	0.976	< 0.001
$VOI_{man-grad+}$	0.991	< 0.001	0.991	< 0.001	0.988	< 0.001
MATV						
$VOI_{man-auto}$	0.309	0.355	0.555	0.077	0.430	0.214
$VOI_{man-grad}$	0.955	< 0.001	0.973	< 0.001	0.806	0.005
$VOI_{man-grad+}$	0.936	< 0.001	1.000	NA	0.964	< 0.001

Spearman's test for correlations was used to calculate significance. The strongest correlation for the three PET scans was found between the VOI_{man} and the VOI_{grad+}, as indicated in gray

VOI volume of interest, *VOI$_{man}$* manually drawn VOI, *VOI$_{auto}$* automatically drawn VOI, *VOI$_{grad}$* VOI based on the gradient between voxels, *VOI$_{grad+}$* VOI$_{grad}$ + necrosis, ^{18}*F-FDG PET-CT* fluorine-18-fluorodeoxyglucose positron emission tomography with computed tomography, *SUV$_{max}$* maximum standardized uptake value, *SUV$_{peak}$* peak standardized uptake value, *SUV$_{mean}$* mean standardized uptake value, *TLG* total lesion glycolysis, *MATV* metabolically active tumor volume, *IQR* interquartile range, *NA* not applicable

Metabolic tumor activity

During neoadjuvant treatment, all five metabolic parameters for the reference VOI_{man} declined between scans 1 and 3 (all $p < 0.05$, Fig. 3, Table 3).

This decline was further explored by calculating the absolute and the percentage difference between the three serial scans. The percentage difference was obtained by dividing the difference between scans by the measured value of the first scan. A significant decline in SUV_{max}, SUV_{peak}, and SUV_{mean} was found between scan 1 vs. scan 2, as well as between scan 1 vs. scan 3. However, no significant decline in SUV_{max}, SUV_{peak}, and SUV_{mean} was found between scan 2 vs. scan 3. The decline in TLG was significant between all serial scans. A significant decline in MATV was found between scan 2 vs. scan 3. The decline in metabolic tumor activity for all parameters except MATV was largest between scan 1 vs. 2, whereas the decline in MATV was largest between scan 2 vs. 3 (Fig. 4, Table 4).

Histopathologic response

Histopathologic response to neoadjuvant treatment varied among the 11 patients, as follows: one grade A (9.1%), one grade B (9.1%), two grade C (18.2%) (totaling to four histopathologic responders (36.4%)), five grade D (45.5%), and two grade E (18.2%) (totaling to 7 non-responders (64.4%)). The histopathologic responders seem to be identifiable by a decline in TLG of > 75% between scans 1 and 3 calculated using the VOI_{man} (Table 5).

To further explore the identification of the histopathologic responders, the difference and percentage difference in TLG between scans 1 and 3 for the four VOI delineation techniques was calculated (Additional file 4: Table S2). A calculated decline in TLG of > 75% using the $VOI_{grad/grad+}$ identified the same histopathologic responders as the VOI_{man}. The VOI_{auto} however failed to identify patient 5 as histopathologic responder. Furthermore, a > 75% decline in TLG was also found with the VOI_{auto} and VOI_{grad} in patients 3 and 4 and with the VOI_{grad+} in patient 4.

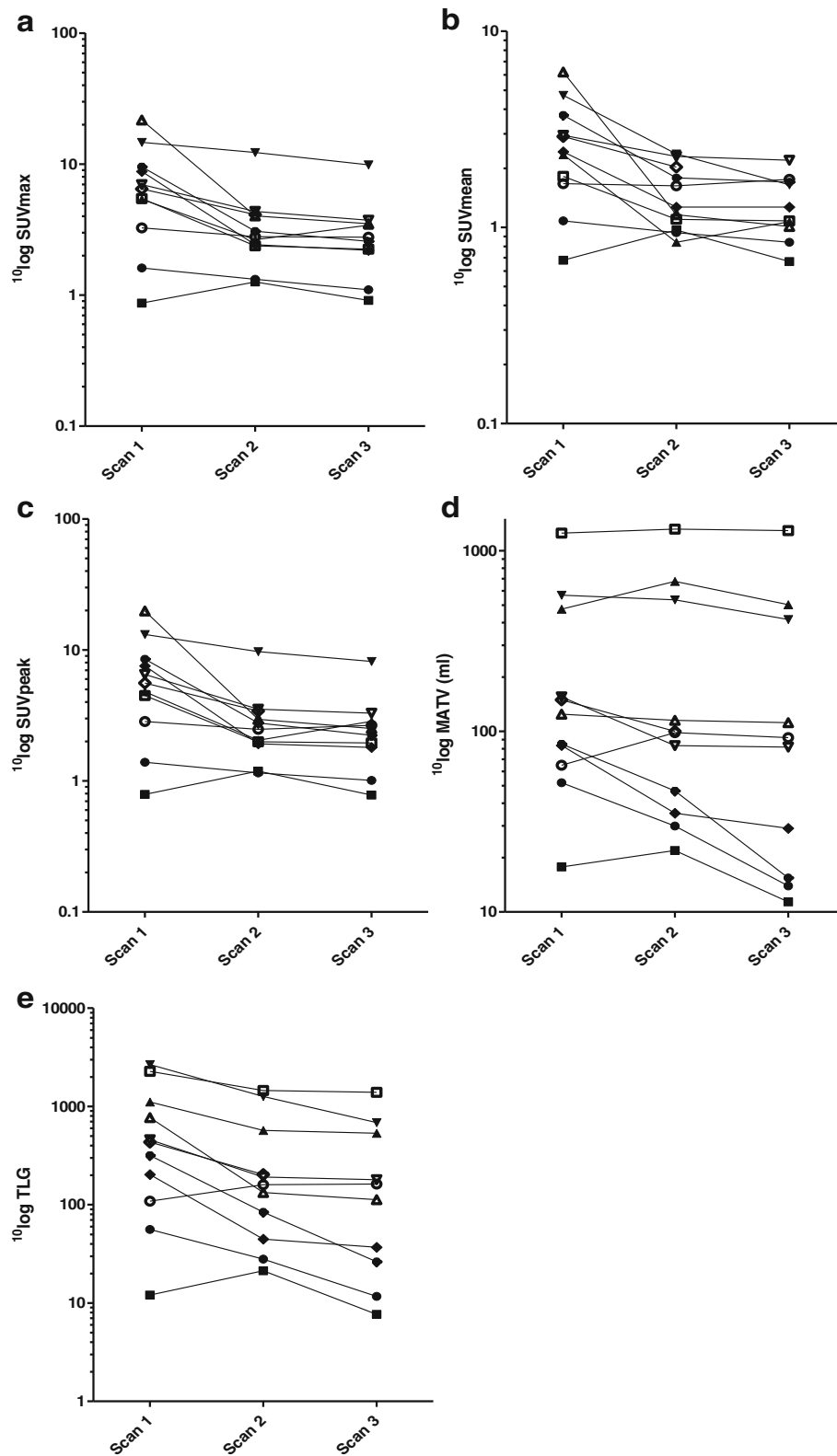Stevenson *et al. EJNMMI Research* (2018) 8:42

Page 6 of 11

Fig. 3 Course in metabolic tumor activity (VOI$_{man}$) during neoadjuvant treatment for each patient individually. The course in metabolic tumor activity for the VOI$_{man}$ during the neoadjuvant treatment for each patient individually for the serial ^{18}F-FDG PET-CT scans. **a** SUV$_{max}$. **b** SUV$_{mean}$. **c** SUV$_{peak}$. **d** Metabolically active tumor-volume (MATV). **e** Total lesion glycolysis (TLG)

Stevenson et al. EJNMMI Research (2018) 8:42

Page 7 of 11

Table 3 Metabolic tumor activity for the VOI_{man} for the serial ^{18}F-FDG PET-CT scans

Parameter	Scan 1	Scan 2	Scan 3	p value
SUV_{max}	6.5 (3.3–9.5)	2.8 (2.4–4.1)	2.7 (1.9–3.6)	0.002
SUV_{peak}	5.6 (2.8–8.5)	2.5 (1.9–3.4)	2.4 (1.6–3.0)	0.001
SUV_{mean}	2.4 (1.7–3.7)	1.3 (1.0–2.0)	1.2 (1.0–1.7)	0.006
TLG	434.8 (108.6–1112.8)	159.9 (44.7–570.9)	137.5 (22.6–572.6)	0.003
MATV (ml)	124.4 (64.8–474.2)	98.3 (35.2–534.2)	87.1 (15.1–437.7)	0.025

Data presented as median (IQR)

VOI volume of interest, VOI_{man} manually drawn VOI, ^{18}F-FDG PET-CT fluorine-18-fluorodeoxyglucose positron emission tomography with computed tomography, SUV_{max} maximum standardized uptake value, SUV_{peak} peak standardized uptake value, SUV_{mean} mean standardized uptake value, *TLG* total lesion glycolysis, *MATV* metabolically active tumor volume, *IQR* interquartile range

Discussion

This study studying four VOI delineation techniques in three consecutive ^{18}F-FDG PET-CT scans per patient demonstrates a significant decline in metabolic tumor activity (VOI_{man}) during the neoadjuvant treatment, consisting of HILP and preoperative EBRT, of locally advanced ESTS. The decline in SUV_{max}, SUV_{peak}, SUV_{mean}, and TLG between scan 1 vs. 2 implies that the HILP accounts for the largest effect on metabolic tumor activity. The MATV seems to be affected most by the EBRT, given the significant decline found between scan 2 vs. 3.

In search of a uniform and reproducible way to calculate changes in metabolic tumor activity in these upfront highly heterogeneous tumors, the use of four different VOI delineation techniques was studied. The VOI_{man} (defined as reference VOI) is the only delineation technique in which the entire tumor is encompassed independently of the amount of necrosis present in the tumor. Therefore, the VOI_{man} delineation technique seems to be most reliable when used for calculating the metabolic tumor activity. However, the VOI_{man} delineation technique is time-consuming, making it unfit for implementation into daily practice. A high correlation, acceptable level of agreement, and comparable ranking was found between the VOI_{man} and the VOI_{grad+} delineation techniques. The differences in ranking between the four VOI delineation techniques are best explained by the high amount of necrosis present in these tumors, as tumor necrosis did not meet the margin thresholds of the VOI_{auto} and VOI_{grad}. To obtain the VOI_{grad+}, the necrosis was manually included, and therefore, the ranking of patients was comparable to the ranking according to the VOI_{man}.

Thus, the VOI_{grad+} delineation technique seems to be a reliable and reproducible technique for the delineation of heterogeneous tumors as ESTS. Further studies including larger patient cohorts in various solid tumor types are necessary for the validation and reproducibility of the various VOI delineation techniques. This study, however, demonstrates that the applied VOI delineation technique is important to consider because we found that assessment of response based on metabolic parameters derived from different VOIs may differ across subjects.

The metabolic tumor changes during neoadjuvant treatment between scan 1 vs. scan 3 were analyzed and compared with the corresponding histopathologic tumor response. Out of the five metabolic parameters tested, TLG seemed to identify the histopathologic responders most reliably (> 75% decrease in TLG between scan 1 and scan 3) when using the VOI_{man} delineation technique. Using the 75% decrease in TLG as a cutoff value was derived empirically from the data, used as example, and to obtain pilot data for using and comparing these techniques. When compared with the VOI_{man} delineation technique, the VOI_{grad+} technique identified the same histopathologic responders with only one additional patient. It seems that these two delineation techniques most reliably identify histopathologic responders, because they include tumor necrosis. The difference in performance of the VOI_{man} and VOI_{grad+} delineation techniques in identifying histopathologic responders is very subtle. However, the VOI_{grad+} delineation technique was found to be easier in use and is considerably less time-consuming than the VOI_{man} technique, making it more suitable for implementation into daily practice. The VOI delineation techniques and the TLG cutoff value need confirmation in larger patient cohorts.

During the last years, the predictive value of ^{18}F-FDG PET-CT scans in staging and monitoring treatment response during neoadjuvant treatment has been established for various solid tumors (including metastatic colorectal cancer and non-small cell lung cancer [23, 27–29]. Therefore, further ESTS studies in which metabolic tumor activity, e.g., > 75% decrease in TLG with VOI_{man} and/ or VOI_{grad+}, is explored as predictor for monitoring therapy response, for histopathologic findings, and for oncological outcome are warranted. The identification of reproducible and reliable VOI delineation techniques, as well as the identification of robust PET parameters for the interpretation of changes in metabolic tumor activity, is relevant because this will enable clinicians to shorten delineation time and to compare
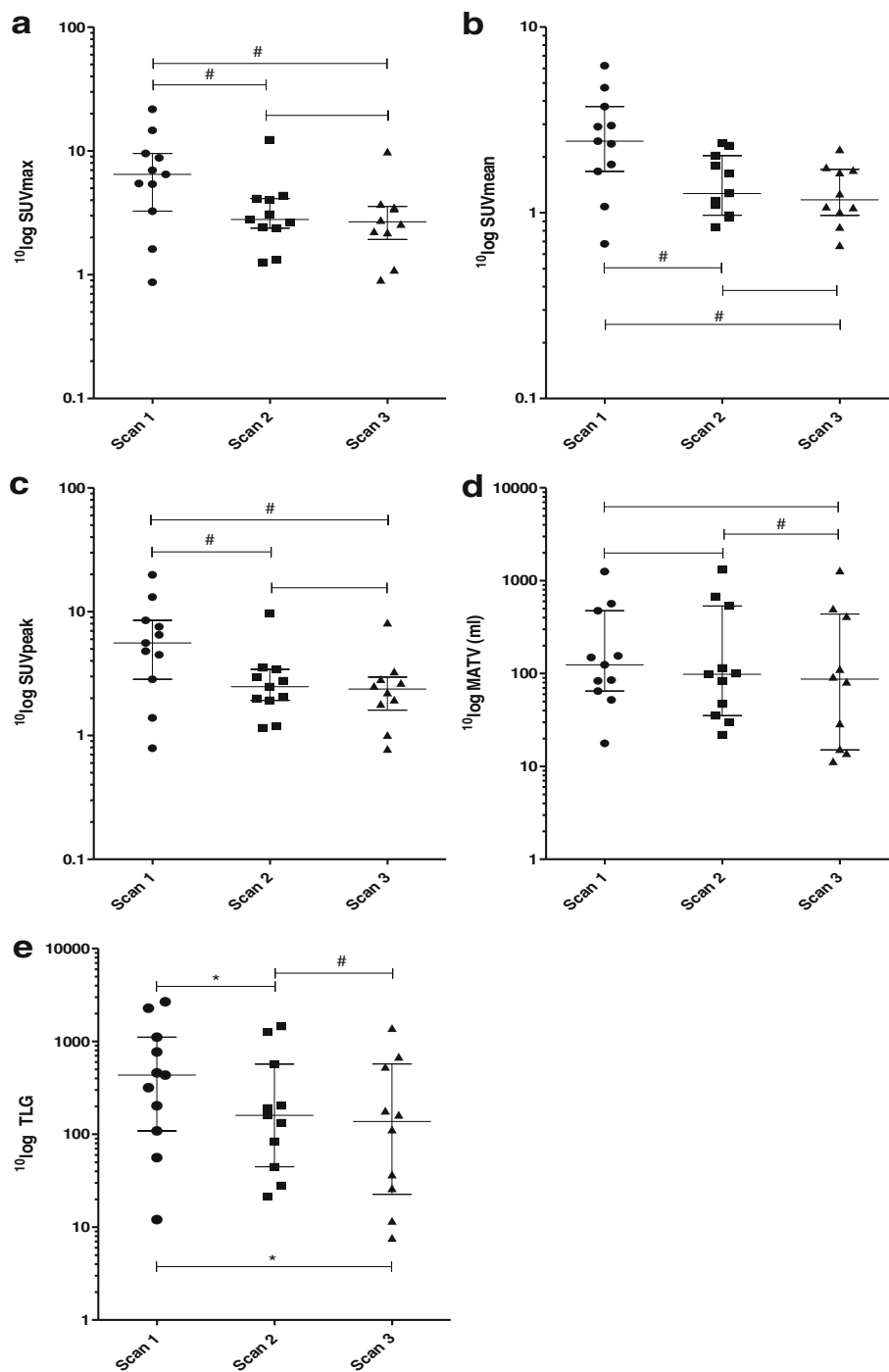
Stevenson *et al. EJNMMI Research* (2018) 8:42

Page 8 of 11

Fig. 4 Changes in metabolic tumor activity (VOl$_{man}$) during neoadjuvant treatment for the serial ^{18}F-FDG PET-CT scans. Changes in metabolic tumor activity for the VOl$_{man}$ during the neoadjuvant treatment for the serial ^{18}F-FDG PET-CT scans. Median and interquartile ranges are indicated. **a** SUV$_{max}$. **b** SUV$_{mean}$. **c** SUV$_{peak}$. **d** Metabolically active tumor volume (MATV). **e** Total lesion glycolysis (TLG). *$p < 0.05$; #$p < 0.01$

results between observers, patients, and centers for ESTS and for other solid tumor types.

This study has some limitations, such as the retrospective character and the small patient population of the study. Only 11 patients were included; however, all

patients but one underwent all three ^{18}F-FDG PET-CT scans, and therefore, it was possible to establish the changes in metabolic tumor activity during the neoadjuvant treatment in all patients. Possibly, the interpretation of the third PET scan is biased by local inflammatory

Stevenson et al. EJNMMI Research (2018) 8:42

Page 9 of 11

Table 4 Changes in metabolic tumor activity for the VOI$_{man}$ during the neoadjuvant treatment between the serial ^{18}F-FDG PET-CT scans

Parameter	Scan 1 vs. 2		Scan 2 vs. 3		Scan 1 vs. 3	
	Δ	Δ %	Δ	Δ %	Δ	Δ %
SUV$_{max}$	− 2.6 (− 6.4 to − 0.5)$^{\#}$	− 37.7 (− 67.7 to − 16.4)	− 0.3 (− 0.5 to − 0.1)	− 13.6 (− 17.5 to − 4.7)	− 3.2 (− 6.7 to − 0.5)$^{\#}$	− 41.6 (− 73.5 to − 27.7)
SUV$_{peak}$	−2.8 (− 5.6 to − 0.4)$^{\#}$	− 45.8 (− 67.4 to − 17.3)	− 0.2 (− 0.5 to 0.0)	− 9.3 (− 16.8 to 0.4)	− 2.9 (− 5.9 to − 0.3)$^{\#}$	− 45.1 (− 74.4 to − 21.9)
SUV$_{mean}$	− 0.9 (− 1.9 to − 0.1)$^{\#}$	− 39.3 (− 52.0 to − 13.2)	− 0.1 (− 0.2 to 0.0)	− 4.6 (− 17.2 to 2.3)	− 1.0 (− 2.3 to − 0.2)$^{\#}$	− 44.1 (− 57.2 to − 17.0)
TLG	− 233.6 (− 637.9 to − 28.0)*	− 52.6 (− 73.6 to − 36.3)	− 18.4 (− 57.0 to − 10.6)$^{\#}$	− 16.3 (− 59.7 to − 5.5)	− 285.0 (− 714.7 to − 34.4)*	− 67.5 (− 82.6 to − 38.2)
MATV (ml)	− 22.1 (− 48.2 to 33.5)	− 7.5 (− 44.9 to 23.5)	− 13.2 (− 53.0 to − 5.3)$^{\#}$	− 19.8 (− 49.4 to − 2.7)	− 25.4 (− 70.6 to 27.6)	− 31.2 (− 67.2 to 3.8)

Data presented as median (IQR)
VOI volume of interest, *VOI$_{man}$* manually drawn VOI, *^{18}F-FDG PET-CT* fluorine-18-fluorodeoxyglucose positron emission tomography with computed tomography, *SUV$_{max}$* maximum standardized uptake value, *SUV$_{peak}$* peak standardized uptake value, *SUV$_{mean}$* mean standardized uptake value, *TLG* total lesion glycolysis, *MATV* metabolically active tumor volume, *IQR* interquartile range. *$p < 0.05$; $^{\#}p < 0.01$

changes following the EBRT. These inflammatory changes might partly explain the significantly more pronounced decrease in metabolic tumor activity following the HILP then following the EBRT, as found in the current series. Despite this potential bias due to radiation-induced local inflammatory changes, a decrease in metabolic tumor activity between scans 1 and 3 was found, which theoretically might have been larger without these changes. For the purpose of this study, all data considering the metabolic tumor activity were obtained from an additional analyses of the ^{18}F-FDG PET-CT scans, since these data are not used in routine patient care. Interestingly, the EORTC-STBSG response score [26] could be used to explore the relationship between changes in metabolic tumor activity and histopathologic response. However, the

prognostic value of the STS response score according to the proportion of stainable tumor cells needs further validation [30].

Conclusions

This study identified the VOI$_{grad+}$ delineation technique as most reliable considering reproducibility when compared with the other delineation techniques during the multimodality neoadjuvant treatment of locally advanced ESTS. Moreover, the VOI$_{grad+}$ delineation technique was considerably less time-consuming to perform when compared to the VOI$_{man}$ technique, potentially resulting in easier implementation in clinical practice. A significant decline in metabolic tumor activity during the treatment was found. The decrease in metabolic tumor

Table 5 Changes in metabolic tumor activity for the VOI$_{man}$ during the neoadjuvant treatment between ^{18}F-FDG PET-CT scans 1 and 3, combined with the corresponding histopathologic tumor response for each patient

Patient number	Scan 1 vs. 3 SUV$_{max}$		Scan 1 vs. 3 SUV$_{peak}$		Scan 1 vs. 3 SUV$_{mean}$		Scan 1 vs. 3 TLG		Scan 1 vs. 3 MATV (ml)		EORTC-STBSG response grade [26]
	Δ	Δ %	Δ	Δ %	Δ	Δ %	Δ	Δ %	Δ	Δ %	
1	− 0.5	− 32.0	− 0.4	− 27.0	− 0.2	− 22.4	− 44.4	− 79.2	− 38.0	− 73.1	C
2	0.1	5.2	− 0.0	− 1.7	0.0	− 0.7	− 4.4	− 36.3	− 6.4	− 35.8	D
3	− 2.0	− 36.6	− 2.0	− 40.7	− 1.3	− 54.6	− 578.1	− 51.9	27.9	5.9	D
4	− 4.8	− 32.7	− 4.9	− 37.6	− 3.1	− 65.1	− 1986.4	− 74.3	− 150.4	− 26.5	D
5	− 6.6	− 75.0	− 5.7	− 76.2	− 1.2	− 47.6	− 165.7	− 81.8	− 54.4	− 65.2	A
6	− 0.5	− 15.0	− 0.2	− 6.4	0.1	5.0	53.8	49.5	27.5	42.4	E
7	− 3.2	− 59.0	− 2.5	− 56.5	− 0.7	− 40.6	− 883.9	− 38.8	38.6	3.1	D
8	− 18.2	− 83.8	− 17.3	− 87.3	− 5.2	− 83.7	− 658.4	− 85.4	− 12.8	− 10.3	B
9	− 3.2	− 46.5	− 3.2	− 49.5	− 0.8	− 25.5	− 278.7	− 60.7	− 73.5	− 47.3	D
10	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	E
11	− 7.0	− 73.1	− 6.3	− 73.8	− 2.0	− 54.4	− 291.3	− 91.7	− 69.6	− 81.9	C

Histopathologic responders are indicated in gray. A percentage difference of > 75% in TLG seemed to identify the histopathologic responders; these values were encircled
^{18}F-FDG PET-CT fluorine-18-fluorodeoxyglucose positron emission tomography with computed tomography, *SUV$_{max}$* maximum standardized uptake value, *SUV$_{peak}$* peak standardized uptake value, *SUV$_{mean}$* mean standardized uptake value, *TLG* total lesion glycolysis, *MATV* metabolically active tumor volume, *EORTC-STBSG* European Organization for Research and Treatment of Cancer-Soft Tissue and Bone Sarcoma Group

Stevenson *et al. EJNMMI Research* (2018) 8:42

Page 10 of 11

activity was significantly more pronounced after HILP than after preoperative radiotherapy. TLG seems promising, but warrants further confirmation, as predictor for histopathologic response in ESTS. Further studies in larger ESTS patient cohorts in which the investigated metabolic parameters and VOI delineation techniques are confirmed and validated as predictors for monitoring treatment response, for histopathologic response, and for oncological outcome are warranted, as this will result in an increase in the clinical applicability of metabolic tumor activity assessments in longitudinal sarcoma ^{18}F-FDG PET-CT studies.

Additional files

Additional file 1: Supplemental Methods. (DOCX 17 kb)

Additional file 2: Figure S1. Bland-Altman plots showing the level of agreement between the VOI$_{man}$ and the VOI$_{auto/grad/grad+}$ for the serial ^{18}F-FDG PET-CT scans for A SUV$_{mean}$, B total lesion glycolysis (TLG), and C metabolically active tumor-volume (MATV). (PDF 529 kb)

Additional file 3: Table S1. Ranking of patients for SUV$_{mean}$, TLG, and MATV according to the four VOI delineation techniques. (DOCX 26 kb)

Additional file 4: Table S2. Changes in TLG according to the four VOI delineation techniques between ^{18}F-FDG PET-CT scan 1 and scan 3, combined with the corresponding histopathologic tumor response for each patient. (DOCX 22 kb)

Abbreviations
^{18}F-FDG PET-CT: Fluorine-18-fluorodeoxyglucose positron emission tomography with computed tomography; EBRT: External beam radiotherapy; EORTC-STBSG: European Organization for Research and Treatment of Cancer-Soft Tissue and Bone Sarcoma Group; ESTS: Extremity soft tissue sarcoma; HILP: Hyperthermic isolated limb perfusion; IQR: Interquartile range; IRB: Institutional Review Board; MATV: Metabolically active tumor volume; STS: Soft tissue sarcoma; SUV$_{max}$: Maximum standardized uptake value; SUV$_{mean}$: Mean standardized uptake value; SUV$_{peak}$: Peak standardized uptake value; TLG: Total lesion glycolysis; VOI: Volume of interest

Availability of data and materials
The data supporting our finding are available upon request.

Authors' contributions
MS, LB, HH, RB, and AB contributed to the study concepts. MS, LB, HH, RB, and AB contributed to the study design. MS, AS, RB, and AB contributed to the data acquisition. MS, AS, RB, and AB contributed to the quality control of data. MS, RB, and AB contributed to the data analyses and interpretation. MS, RB, and AB contributed to the statistical analysis. MS, LB, RB, and AB contributed to the manuscript preparation. MS, LB, HH, AS, RB, and AB contributed to the manuscript editing. MS, LB, HH, AS, RB, and AB contributed to the manuscript review. All authors read and approved the final manuscript.

Ethics approval and consent to participate
This study has been approved by the Institutional Review Board (IRB), and the need for written informed consent was waived (IRB case number 2016.984).

Competing interests
The authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
^{1}Department of Surgical Oncology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. ^{2}Department of Pathology and Medical Biology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. ^{3}Department of Nuclear Medicine and Molecular Imaging, University Medical Center Groningen, University of Groningen, Hanzeplein 1, PO Box 30.001, 9700 RB Groningen, The Netherlands. ^{4}Department of Radiology and Nuclear Medicine, VU University Medical Center, Amsterdam, The Netherlands.

References
1. Soft tissue sarcoma incidence, Nederlandse Kankerregistratie, beheerd door IKNL © [March] 2018. Available at: www.cijfersoverkanker.nl. Accessed March, 2018.
2. Soft tissue sarcoma deaths, Nederlandse Kankerregistratie, beheerd door IKNL© [March] 2018. Available at: www.cijfersoverkanker.nl. Accessed March, 2018.
3. Morrison BA. Soft tissue sarcomas of the extremities. Proc (Bayl Univ Med Cent). 2003;16(3):285–90.
4. Hoekstra HJ, Haas RLM, Verhoef C, Suurmeijer AJH, van Rijswijk CSP, Bongers BGH, et al. Adherence to guidelines for adult (non-GIST) soft tissue sarcoma in the Netherlands: a plea for dedicated sarcoma centers. Ann Surg Oncol. 2017;24(11):3279–88.
5. Eggermont AM, Schraffordt Koops H, Klausner JM, Kroon BB, Schlag PM, Lienard D, et al. Isolated limb perfusion with tumor necrosis factor and melphalan for limb salvage in 186 patients with locally advanced soft tissue extremity sarcomas. The cumulative multicenter European experience. Ann Surg. 1996;224(6):756–64. discussion 764-5
6. Eggermont AM, Schraffordt Koops H, Lienard D, Kroon BB, van Geel AN, Hoekstra HJ, et al. Isolated limb perfusion with high-dose tumor necrosis factor-alpha in combination with interferon-gamma and melphalan for nonresectable extremity soft tissue sarcomas: a multicenter trial. J Clin Oncol. 1996;14(10):2653–65.
7. Hoven-Gondrie ML, Bastiaannet E, van Ginkel RJ, Pras EB, Suurmeijer A, Hoekstra HJ. Limb perfusion in soft tissue sarcomas: twenty years of experience. Ned Tijdschr Geneeskd. 2013;157(30):A6148.
8. Hoven-Gondrie ML, Bastiaannet E, van Ginkel RJ, Suurmeijer AJ, Hoekstra HJ. TNF dose reduction and shortening of duration of isolated limb perfusion for locally advanced soft tissue sarcoma of the extremities is safe and effective in terms of long-term patient outcome. J Surg Oncol. 2011;103(7):648–55.
9. Bhangu A, Broom L, Nepogodiev D, Gourevitch D, Desai A. Outcomes of isolated limb perfusion in the treatment of extremity soft tissue sarcoma: a systematic review. Eur J Surg Oncol. 2013;39(4):311–9.
10. Adjuvant chemotherapy for localised resectable soft tissue sarcoma in adults. Sarcoma Meta-analysis Collaboration (SMAC). Cochrane Database Syst Rev. 2000;(2):CD001419.
11. Le Cesne A, Ouali M, Leahy MG, Santoro A, Hoekstra HJ, Hohenberger P, et al. Doxorubicin-based adjuvant chemotherapy in soft tissue sarcoma: pooled analysis of two STBSG-EORTC phase III clinical trials. Ann Oncol. 2014;25(12):2425–32.
12. Saponara M, Stacchiotti S, Casali PG, Gronchi A. (Neo)adjuvant treatment in localised soft tissue sarcoma: the unsolved affair. Eur J Cancer. 2017;70:1–11.
13. van Ginkel RJ, Hoekstra HJ, Pruim J, Nieweg OE, Molenaar WM, Paans AM, et al. FDG-PET to evaluate response to hyperthermic isolated limb perfusion for locally advanced soft-tissue sarcoma. J Nucl Med. 1996;37(6):984–90.
14. Chen L, Wu X, Ma X, Guo L, Zhu C, Li Q. Prognostic value of 18F-FDG PET-CT-based functional parameters in patients with soft tissue sarcoma: a meta-analysis. Medicine (Baltimore). 2017;96(6):e5913.
15. Casey DL, Wexler LH, Fox JJ, Dharmarajan KV, Schoder H, Price AN, et al. Predicting outcome in patients with rhabdomyosarcoma: role of [(18)f]fluorodeoxyglucose positron emission tomography. Int J Radiat Oncol Biol Phys. 2014;90(5):1136–42.
16. Nishiyama Y, Tateishi U, Kawai A, Chuman H, Nakatani F, Miyake M, et al. Prediction of treatment outcomes in patients with chest wall sarcoma: evaluation with PET/CT. Jpn J Clin Oncol. 2012;42(10):912–8.

Stevenson *et al. EJNMMI Research* (2018) 8:42

Page 11 of 11

17. 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(9):1519–27.

18. Lodge MA, Chaudhry MA, Wahl RL. Noise considerations for PET quantification using maximum and peak standardized uptake value. J Nucl Med. 2012;53(7):1041–7.

19. Stevenson MG, Seinen JM, Pras E, Brouwers AH, van Ginkel RJ, van Leeuwen BL, et al. Hyperthermic isolated limb perfusion, preoperative radiotherapy and surgery (PRS) a new limb saving treatment strategy for locally advanced sarcomas. J Surg Oncol. 2018; https://doi.org/10.1002/jso.25008.

20. Boellaard R, O'Doherty MJ, Weber WA, Mottaghy FM, Lonsdale MN, Stroobants SG, et al. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. Eur J Nucl Med Mol Imaging. 2010;37(1):181–200.

21. Boellaard R, Delgado-Bolton R, Oyen WJ, Giammarile F, Tatsch K, Eschner W, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. Eur J Nucl Med Mol Imaging. 2015;42(2):328–54.

22. Frings V, van Velden FH, Velasquez LM, Hayes W, van de Ven PM, Hoekstra OS, et al. Repeatability of metabolically active tumor volume measurements with FDG PET/CT in advanced gastrointestinal malignancies: a multicenter study. Radiology. 2014;273(2):539–48.

23. Kramer GM, Frings V, Hoetjes N, Hoekstra OS, Smit EF, de Langen AJ, et al. Repeatability of quantitative whole-body 18F-FDG PET/CT uptake measures as function of uptake interval and lesion selection in non-small cell lung cancer patients. J Nucl Med. 2016;57(9):1343–9.

24. Boellaard R. Quantitative oncology molecular analysis suite: ACCURATE. SNMMI June 23-26, 2018.

25. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;1(8476):307–10.

26. Wardelmann E, Haas RL, Bovee JV, Terrier P, Lazar A, Messiou C, et al. Evaluation of response after neoadjuvant treatment in soft tissue sarcomas; the European Organization for Research and Treatment of Cancer-Soft Tissue and Bone Sarcoma Group (EORTC-STBSG) recommendations for pathological examination and reporting. Eur J Cancer. 2016;53:84–95.

27. van Helden EJ, Hoekstra OS, Boellaard R, Roth C, Mulder ER, Verheul HM, et al. Early 18F-FDG PET/CT evaluation shows heterogeneous metabolic responses to anti-EGFR therapy in patients with metastatic colorectal cancer. PLoS One. 2016;11(5):e0155178.

28. Soussan M, Cyrta J, Pouliquen C, Chouahnia K, Orlhac F, Martinod E, et al. Fluorine 18 fluorodeoxyglucose PET/CT volume-based indices in locally advanced non-small cell lung cancer: prediction of residual viable tumor after induction chemotherapy. Radiology. 2014;272(3):875–84.

29. Chen HH, Chiu NT, Su WC, Guo HR, Lee BF. Prognostic value of whole-body total lesion glycolysis at pretreatment FDG PET/CT in non-small cell lung cancer. Radiology. 2012;264(2):559–66.

30. Schaefer IM, Hornick JL, Barysauskas CM, Raut CP, Patel SA, Royce TJ, et al. Histologic appearance after preoperative radiation therapy for soft tissue sarcoma: assessment of the European Organization for Research and Treatment of Cancer-Soft Tissue and Bone Sarcoma Group Response Score. Int J Radiat Oncol Biol Phys. 2017;98(2):375–83.