Performance of automatic image segmentation algorithms for calculating total lesion glycolysis for early response monitoring in non-small cell lung cancer patients during concomitant chemoradiotherapy

Willem Grootjans[1,2*], Edwin A. Usmanij[1*], Wim J.G. Oyen[1,3], Erik H.F.M. van der Heijden[4], Eric P. Visser[1], Dimitris Visvikis[5], Mathieu Hatt[5], Johan Bussink[6], Lioe-Fee de Geus-Oei[1,2,7]

[*]*Willem Grootjans and Edwin A. Usmanij contributed equally to this work*

[1]Department of Radiology and Nuclear Medicine, Radboud University Medical Center, Nijmegen, The Netherlands.

[2]Department of Radiology, Leiden University Medical Center, Leiden, The Netherlands.

[3]The Institute of Cancer Research and The Royal Marsden NHS Foundation Trust, London, UK

[4]Department of Pulmonary Medicine, Radboud University Medical Center, Nijmegen, The Netherlands.

[5]INSERM, UMR 1101, Laboratoire de Traitement de l'information Médicale (LaTIM), Brest, France

[6]Department of Radiation Oncology, Radboud University Medical Center, Nijmegen, The Netherlands.

[7]Biomedical Photonic Imaging Group, MIRA Institute, University of Twente, Enschede, The Netherlands.

Institution where the work was performed:

1. Department of Radiology and Nuclear Medicine, Radboud university medical center, Nijmegen, The Netherlands

Correspondence: W. Grootjans M.Sc.

Department of Radiology

Leiden University Medical Center

P.O. Box 9600, 2300 RC Leiden, The Netherlands

E-mail: W.Grootjans@lumc.nl

Manuscript statistics:

Number of pages: 22

Number of figures: 2

Number of tables: 3

Running head: FDG-PET in early NSCLC response assessment

Keywords:

Early response monitoring; Non-small cell lung cancer; concomitant chemoradiochemotherapy; ^{18}F-FDG PET/CT, total lesion glycolysis, automatic image segmentation

Abstract

Background and purpose:

This study evaluated the use of total lesion glycolysis (TLG) determined by different automatic segmentation algorithms, for early response monitoring in non-small cell lung cancer (NSCLC) patients during concomitant chemoradiotherapy.

Materials and Methods:

Twenty-seven patients with locally advanced NSCLC treated with concomitant chemoradiotherapy underwent ^{18}F-fluorodeoxyglucose (FDG) PET/CT imaging before and in the second week of treatment. Segmentation of the primary tumours and lymph nodes was performed using fixed threshold segmentation at (i) 40% SUV_{max} (T40), (ii) 50% SUV_{max} (T50), (iii) relative-threshold-level (RTL), (iv) signal-to-background ratio (SBR), and (v) fuzzy locally adaptive Bayesian (FLAB) segmentation. Association of primary tumour TLG (TLG_T), lymph node TLG (TLG_{LN}), summed TLG (TLG_S=TLG_T+TLG_{LN}), and relative TLG decrease (ΔTLG) with overall-survival (OS) and progression-free survival (PFS) was determined using univariate Cox regression models.

Results:

Pre-treatment TLG_T was predictive for PFS and OS, irrespective of segmentation method used. Inclusion of TLG_{LN} improved early response assessment, with pre-treatment TLG_Smore stronglyassociated with PFS and OS than $TLG_T$for all segmentation algorithms. This was also the case for ΔTLG_S, which was significantly associated with PFS and OS, with exception of RTL and T40.

Conclusions:

ΔTLG_S was significantly associated with PFS and OS, except for RTL and T40. Inclusion of TLG_{LN} improves early treatment response monitoring during concomitant chemoradiotherapy with FDG-PET.

Introduction

Non-small cell lung (NSCLC) cancer remains a disease with a generally poor prognosis [1]. At the time of diagnosis, one third of patients with NSCLC presents with locally advanced non-metastatic disease [1]. For these patients, radiotherapy in combination with chemotherapy is the accepted standard of care. With the aim of improving patient outcome, combined and intensified treatment approaches are increasingly being investigated. However, not all patients equally benefit from these treatment approaches and rational selection of available treatment options in a personalized medicine framework is required[2].

Positron emission tomography (PET) in combination with X-ray computed tomography (CT) with the glucose analogue ^{18}F-fluorodeoxyglucose (FDG) has proven to be a valuable tool to personalize treatment for this patient group[2]. Firstly, incorporation of FDG-PET images into the radiotherapy planning algorithm improves definition of gross tumour volume (GTV)[3-5] and might facilitate the concept of selective nodal irradiation[6]. Secondly, it has been shown that FDG-PET can identify areas that are at risk of local relapse[7, 8], permitting to use the concept of molecular imaging-based dose painting[9]. Thirdly, several studies emphasize the ability of FDG-PET to monitor therapy response at an early treatment stage using quantitative PET indices [10-14]. Early response monitoring during treatment can facilitate clinical decision-making and improve patient management through avoidance of unnecessary side effects and costs of ineffective treatment.

However, employing the concept of FDG-PET-guided treatment decisions requires robust and standardized methods to derive these quantitative indices from PET images. Particularly, the strong dependence of most image-derived response indices on quantification of these volumes emphasizes the need for standardized and consistent determination of lesion volume in PET images. In this regard, there has been a widespread interest in the development for automated segmentation algorithms for PET.

Over the years, there has been a rapid growth of segmentation algorithms for PET reported in literature [15], an event which is also referred to as '*yapetism*' ("yet another PET segmentation method") [16]. Difficulties encountered by these algorithms for automatic lesion segmentation in PET images are local contrast variations due to heterogeneous FDG uptake in the lesion, adjacent FDG-avid anatomy and lymph nodes, and relatively high noise content of PET images, often rendering the task of automatic lesion delineation challenging [15]. This becomes even more difficult when automatic segmentation has to be performed on low contrast interim and end-of-treatment PET images, where radiotracer uptake can be considerably reduced due to therapy effects. However, up until this day there is no standardized method for automatically determining lesion volume on PET images and many studies consider different segmentation algorithms for this purpose [17, 18]. The purpose of this study was to evaluate this clinical applicability and performance of several established segmentation algorithms for generating plausible segmentation volumes that can be applied specifically to predict therapy response during treatment for patients with locally advanced stage IIIA or IIIB NSCLC treated with concomitant chemoradiotherapy. The predictive value of total lesion glycolysis (TLG), as determined by these different algorithms, for early response assessment during concomitant chemoradiotherapy was evaluated.

Materials and methods

Patients

A total of 27 patients with newly diagnosed NSCLC stage IIIA or stage IIIB were prospectively included in this study, as described before [10]. Patients were treated with concomitant radiotherapy and chemotherapy. This study was approved by the institutional review board (IRB) of the Radboud university medical center. Written informed consent was obtained from every patient. Patient characteristics are summarized in Table 1.

Treatment and follow-up

Intensity modulated radiotherapy (IMRT) was performed (10 MV photons), consisting of 33 fractions of 2 Gy (5 fractions a week for 6 week and 3 days) resulting in a total dose of 66 Gy on the primary tumor and affected lymph nodes (i.e. pathologically proven or FDG-avid lymph nodes). Chemotherapy consisted of two cycles of cisplatin 50 mg/m^2 intravenously (day 1, 8, 22, and 29) and etoposide 100 mg/m^2 intravenously (day 1-3, and day 22-24). Median overall treatment time was 45 days (range 43-48 days). Patients with progressive disease during follow-up received palliative treatment. Follow-up during and after treatment consisted of clinical examination at regular intervals. When residual or recurrent disease was suspected, chest X-ray and chest CT-scans were performed. For each patient, sequential FDG-PET/CT imaging was performed before and during treatment. The pre-treatment scan was obtained before treatment (median 11 days, range 1-28 days) while interim FDG-PET imaging was performed in the second week during concomitant treatment (median 14 days, range 13-16 days), always before the second cycle of chemotherapy after 20 Gy radiotherapy. According to the treatment protocol all patients started with radiotherapy at the first day of the first cycle of chemotherapy, i.e. no neo-adjuvant treatment was applied.

Patient preparation and FDG PET imaging

Imaging was performed using a hybrid Biograph Duo PET/CT scanner (Siemens Medical Solution, Knoxville, TN, USA). The PET scanner was accredited by the Research 4 Life (EARL) initiative for quantitative FDG-PET/CT studies [19]. Before image acquisition, patients fasted for at least six hours and blood glucose levels were lower than 8.2 mmol·L^{-1} in all patients. The amount of activity administered to the patient was adjusted to the patient's weight and was 3.45 MBq·kg^{-1}. Details regarding the PET acquisition protocol are summarized in table 1.For the purpose of attenuation correction and

anatomical reference, a low dose (LD) CT scan was acquired with a reference tube current time product of 40 mA·s. LDCT scans were acquired during timed unforced expiration breath-hold. Modulation of X-ray tube current was performed using CARE Dose 4D. Reconstruction of PET images was performed with a 2D ordered subset expectation maximization (2DOSEM) algorithm using a matrix size of 128×128, 4 iterations and 16 subsets. Post reconstruction filtering was performed using a three-dimensional Gaussian filter kernel with a full width half maximum of 5 mm.

Image segmentation

The primary tumour and FDG positive lymph nodes were delineated on the pre-treatment and interim PET images. Firstly, delineation was performed using a fixed threshold region growing segmentation at 40% (T40) and 50% (T50) of the maximum standardized uptake value (SUV_{max}) value. Furthermore, adaptive threshold algorithms were used for image segmentation. These included the iterative relative-threshold-level (RTL)[20] and signal-to-background ratio (SBR)[21] approach. For the SBR method, the background for segmentation of the primary tumour was defined by placing a volume of interest (VOI) in parenchyma of the contra-lateral lung. For lymph node segmentation, the background was defined by placing a VOI near the aortic root in the mediastinum. The seed-point for the T40, T50, RTL, and SBR segmentation was the SUV_{max} voxel of either the primary tumour ($SUV_{T,max}$) or the corresponding lymph nodes ($SUV_{LN,max}$). The threshold-based segmentations were performed using the Inveon Research Workplace (IRW) 4.1 Software (Preclinical Solutions, Siemens Medical Solutions USA, Knoxville Tennessee, USA). In addition to threshold-based segmentation, image segmentation was performed using the fuzzy locally adaptive Bayesian (FLAB) algorithm [22]. Segmentation with this algorithm is performed using custom in-house developed software (ImageD, LaTIM Université de Bretagne Occidentale, Brest, France). The number of classes for segmentation was limited to two and parameters were automatically determined by the software.

Image analysis

The TLG of the primary tumour (TLG_T), defined as the product of the mean tumour FDG uptake ($SUV_{T,mean}$) and metabolic tumour volume (MTV), was calculated on the pre-treatment and interim PET images. Similarly, lymph node TLG (TLG_{LN}) was defined as the mean uptake of the lymph nodes ($SUV_{LN,mean}$) and the corresponding metabolic volumes of the lymph nodes. Furthermore, a summed TLG ($TLG_S=TLG_T+TLG_{LN}$ was calculated. Evaluation of therapy response was performed by calculating the fractional decrease in TLG between pre-teatment and interim PET images (ΔTLG). Segmentation performance of the different algorithms was evaluated through visual assessment by a nuclear medicine physician experienced in thoracic imaging. Segmentation failures were visually identified andwere defined as the propagation of segmentation into other anatomical structures, or premature termination of the algorithm resulting in only partial segmentation of the primary tumour and lymph nodes. Lesions that could not be properly segmented according to these criteria were omitted from the analysis. In addition, similarity between MTVs obtained with different segmentation algorithms was quantified by calculating the spatial overlap using a generalization of the Jaccard index (JI), as described in equation 1.

$$JI(A, B, \dots, n) = \frac{|A \cap B \cap \dots \cap n|}{|A \cup B \cup \dots \cup n|} \qquad 1$$

Here the numerator $|A \cap B \cap \dots \cap n|$ denotes the intersection between segmented volumes (in this study, five in total), while the denominator $|A \cup B \cup \dots \cup n|$ represents the union of the segmented volumes. Perfect spatial overlap is indicated by a value for the JI of 1.0, whilst a value of 0 indicates no spatial overlap of the volumes.

Clinical outcome and statistical analysis

Statistical analysis was performed using SPSS Statistics 20 (IBM, Armonk, New York, USA) and GraphPad Prism, version 4.0c (GraphPad Software Inc, La Jolla, California, USA). Patient outcome data for time to progression was defined as the interval between the start of treatment and the date of documented disease progression as confirmed by imaging or biopsy. If a patient was progression free at the closeout-date (January, 2015), time to progression was censored to that date. PFS was measured from the date of treatment start to the date of documented disease progression. Similarly, if patients were still alive at the closeout-date, patients were censored for OS. The predictive value of pre-treatment TLG and ΔTLG were determined for different segmentation algorithms using univariate Cox regression models. Correlation between the MTVs and TLGs of different segmentation algorithms was calculated using Spearman rank correlation. Statistical significance was defined for $p<0.05$.

Results

The median follow-up time for this patient population was 23.4 months (range 3.5 – 61.9). During follow-up eighteen patients died, all related to cancer progression. Three patients were lost during follow-up. A total of twenty patients developed recurrent disease during follow-up; seven patients developed progression of local disease, while metastases were seen in 13 patients. Median time to disease progression was 21 months. PFS after study-baseline at 1 year was 63% (17 out of 27).

Of the 27 patients, 25 had a visible primary tumour. For the other two patients, there was no radiological evidence for a primary tumour (i.e. cT_0). The smallest pre-treatment MTVs were obtained when segmentation was performed with the T50 segmentation algorithm (24.7 ± 30.8 mL), compared to

T40 (34.8 ± 39.1 mL), SBR (38.7 ± 42.4 mL) RTL (30.7 ± 34.0 mL). Segmentation with FLAB resulted in the largest MTVs (42.3 ± 42.1 mL). Interim PET MTVs showed similar trends with the smallest MTV for the T50 (20.8 ± 32.6 mL) method, followed by T40 (31.6 ± 44.6 mL), RTL (26.8 ± 37.0 mL), SBR (36.2 ± 50.2 mL) and FLAB (37.6 ± 48.2 mL). For the pre-treatment PET images, excellent correlation was found between delineation methods for MTV (range ρ 0.97 – 1.0, $p < 0.0005$), TLG_T (range ρ 0.95 – 1.0, $p < 0.0005$) and SUVmean (range ρ 0.98 – 1.0, $p < 0.0005$). Similarly, for in-treatment PET images, an excellent correlation was found between delineation methods for MTV (range ρ 0.93 – 1.0, $p < 0.0005$), TLG_T (range ρ 0.94 – 1.0, $p < 0.0005$) and SUVmean (range ρ 0.94 – 1.0, $p < 0.0005$).

Quantitative and visual analysis of the MTVs obtained with different segmentation algorithms revealed that the MTVs were highly similar regarding shape and spatial overlap. Furthermore, the algorithms revealed a similar trend in $SUV_{T,mean}$ and MTV change between pre-treatment and interim FDG-PET (supplementary data). The T50 volumes were always completely enveloped by the other volumes. The generalized JI for all MTVs on the pre-treatment and interim PET was 0.58 ± 0.13 (range 0.31 – 0.78) and 0.53 ± 0.16 (range 0.20 – 0.86). When the T50 volumes were omitted from the analysis, the mean JI for pre-treatment and interim PET volumes was 0.74±0.12 (range 0.46 – 0.89) and 0.71 ± 0.14 (range 0.47 – 0.96). In tumours with heterogeneous FDG uptake, the T50 algorithm yielded contours that were more erratic and sensitive to discontinuities within the tumour, while FLAB, SBR, RTL, and T40 algorithms would segment patches with FDG-uptake continuously throughout the entire lesion, giving an improved representation of the total volume with FDG-uptake.

In the pre-treatment PET images, FLAB segmentation resulted in one segmentation failure of the primary tumour in one patient due to small size and low contrast. In this patient, all segmentation methods failed to segment the primary tumour in the interim PET images. Furthermore, the T40, T50, RTL, and SBR methods failed to segment the primary tumour in an additional patient that was presenting with a large lesion with extended growth into the central mediastinum on interim PET

images. The FLAB algorithm could segment the primary tumour in this patient and did not show uncontrolled propagation of segmentation into the mediastinal background.

Of the 27 patients, 18 patients had FDG positive lymph nodes on the pre-treatment PET images. Given the smaller volumes and in general lower contrast of mediastinal and hilar lymph nodes, there were considerably more segmentation failures when performing automatic segmentation of lymph nodes. The T40 and RTL algorithms failed to segment 14 and 17 of the 41 lymph nodes on the pre-treatment PET images, respectively. The number of segmentation failures for the T50 and FLAB algorithms in the pre-treatment PET images was 9 and 8, respectively. The SBR algorithm had the fewest segmentation failures, with only 6 lymph nodes segmentation failures in the pre-treatment PET images.

Reduction of lymph node contrast owing to therapy effects on the interim PET images resulted in more segmentation failures for the T40, T50, RTL and FLAB algorithms. Of the 41 lymph nodes in the interim PET, there were 24 segmentation failures for the T40 and RTL algorithms. For the T50 and FLAB method, 12 and 14 lymph node segmentation failures occurred in the interim PET, respectively. Similar to the pre-treatment PET, the SBR algorithm had the fewest segmentation failures, with 5 segmentation failures. Failure of lymph node segmentation was usually due to uncontrolled propagation of the segmentation algorithm in the mediastinal background or primary tumour. Figure 1 depicts the TLG_S on pre-treatment and interim PET images in box whisker plots.

Pre-treatment $SUV_{T,max}$, interim $SUV_{T,max}$, and relative decrease in $SUV_{T,max}$ of the primary tumour was not significantly predictive for PFS and OS. Similarly, pre-treatment $SUV_{T,mean}$, interim $SUV_{T,mean}$, and relative decrease in $SUV_{T,mean}$ of the primary tumour was not significantly predictive for PFS and OS in this cohort. However, pre-treatment TLG_T was significantly associated with PFS and OS for all segmentation methods. The ΔTLG_T was significantly predictive with PFS for all methods except for FLAB. Furthermore, ΔTLG_T was significantly associated with OS for the T50 and SBR methods. Hazard ratios

(HRs) and corresponding 95% confidence interval (CI) of TLG_T and ΔTLG_T in the univariate Cox regression analysis for PFS and OS are summarized in table 2.

Only TLG_{LN} obtained with the SBR and T50 methods was significantly predictive for PFS and OS. Furthermore, inclusion of TLG_{LN} improved early response assessment using PET, with pre-treatment TLG_S more strongly associated with PFS and OS than TLG_T. Figure 2 depicts differences in PFS and OS of two patients with a different lymph node response, as reflected by ΔTLG_S.

The differences in lymph node segmentation performance was reflected in the significance of TLG_S measurements in the univariate analysis. The HRs and corresponding 95% CI of TLG_S and ΔTLG_S in the univariate Cox regression analysis for PFS and OS are summarized in table 3.

Discussion

In this study we showed that TLG is a robust metric to monitor therapy response in patients undergoing concomitant chemoradiotherapy for locally advanced NSCLC. Furthermore, inclusion of TLG_{LN} improves early assessment of treatment response in this patient population. Results of this study are in line with available data in literature and emphasize the role of FDG-PET imaging for early response monitoring NSCLC [10-14]. In particular, TLG outperformed the more traditional $SUV_{T,max}$ and $SUV_{T,mean}$ for predicting PFS and OS. This is probably due to the fact that TLG contains information about disease load as well as metabolic activity of involved lesions.

In general, all segmentation algorithms had a similar performance for segmenting the primary tumour in different anatomical locations. The presence of adjacent anatomical structures (e.g. lymph nodes, mediastinum, liver), did not result in large differences in segmentation performance. Furthermore, lower contrast of the interim PET images resulted in a very limited increase in the number of segmentation failures. Absolute differences in TLG obtained by the different segmentation methods did

not influence its predictive value. . Due to the limited size and lower contrast, there were considerably more lymph node segmentation failures. The number of lymph node segmentation failures increased in the interim PET images owing to further reduction in image contrast. Nevertheless, adequate lymph node segmentation is of importance, with TLG_S having a stronger association with PFS and OS. Out of all the segmentation algorithms, the SBR method demonstrated the lowest number of segmentation failures. The number of FLAB lymph node segmentation failures could be reduced by using a supervised input, with an equal performance to that of the SBR method, which is in line with results from another study [23]. However, in view of standardizing response measurements, such a user dependency should be avoided and we chose only to include the results of automatic FLAB segmentation.

Although the results emphasize that PET could be used for prediction of early treatment response in patients with locally advanced NSCLC treated with concomitant chemoradiotherapy, employment of PET-guided decisions for personalizing treatment was not explored. Particularly, the strong association of pre-treatment TLG with PFS and OS might merit the choice for treatment intensification in patients with a high pre-treatment TLG such as proposed in the PET-boost dose-escalation trial[24]. Furthermore, one might also consider treatment intensification when interim PET images demonstrate a limited decrease in TLG, for instance by dose escalation to metabolically active sub volumes the primary tumour[25, 26], in order to improve loco-regional tumour control. However, standardizing PET-based dose painting approaches is of utmost importance. This is emphasized in a study by Knudtsen et al. where the used PET reconstruction algorithm and choice of segmentation thresholds significantly influenced treatment plans incorporating these dose painting concepts [27]. Although threshold-based segmentation is frequently used for defining sub volumes for dose painting, stability of different algorithms under varying imaging conditions for this purpose has yet to be investigated. Interestingly, studies emphasize that there is a high stability of FDG uptake in tumour areas during the course of the treatment that can be identified on pre-treatment FDG-PET images[28]. Although useful, interim PET

imaging in a molecularly image-guided ART (IGART) setting would still be of great interest to monitor changes during radiotherapy [2]. Indeed, employment of IGART using FDG-PET has been shown to be of potential value, where the GTV is adapted according to interim FDG-PET imaging[2].Furthermore, results from a study conducted by Nygård et al. emphasize that FDG-PET might identify lesion-specific response after a single series of chemotherapy in NSCLC and could be a useful addition to guide and individualize radiotherapy strategy [29]. Although dose redistribution might be useful for improving loco-regional tumour control, systemic disease control is also an important aspect in this patient group [30]. In this setting, interim PET imaging might identify failure of systemic disease control at an early stage (i.e. detection of additional lymph node or distant metastasis), making it possible to adapt treatment accordingly.

A limitation of the current study is that only a small patient cohort was considered. However, the advantage is that TLG measurements using different automatic segmentation algorithms showed consistent results, with most algorithms yielding TLG values that had a similar predictive value in this patient cohort.

This study emphasizes that adequate lymph node segmentation in PET images improves assessment of early treatment response in NSCLC patients treated with concomitant chemoradiotherapy. In this regard, given the relatively ease of implementation and the high number of successful lymph node segmentations, SBR is the method of choice for calculation of TLG in FDG-PET images of patients with locally advanced NSCLC for the purpose of assessment of early treatment response.

Acknowledgements

Willem Grootjans is the recipient of an educational grant from Siemens Healthcare, The Hague, The Netherlands

Conflict of interest

The authors declare no competing interests

Ethical considerations

All procedures performed in this study were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

References

[1] Aupérin A, Le Péchoux C, Rolland E et al. Meta-Analysis of Concomitant Versus Sequential Radiochemotherapy in Locally Advanced Non–Small-Cell Lung Cancer. J Clin Oncol. 2010;28:2181-90.

[2] Grootjans W, de Geus-Oei L-F, Troost EGC, Visser EP, Oyen WJG, Bussink J. PET in the management of locally advanced and metastatic NSCLC. Nat Rev Clin Oncol. 2015;12:395-407.

[3] van Baardwijk A, Bosmans G, Boersma L et al. PET-CT–Based Auto-Contouring in Non–Small-Cell Lung Cancer Correlates With Pathology and Reduces Interobserver Variability in the Delineation of the Primary Tumor and Involved Nodal Volumes. Int J Radiat Oncol Biol Phys. 2007;68:771-8.

[4] De Ruysscher D, Nestle U, Jeraj R, MacManus M. PET scans in radiotherapy planning of lung cancer. Lung Cancer. 2012;75:141-5.

[5] Chirindel A, Adebahr S, Schuster D et al. Impact of 4D-^{18}FDG-PET/CT imaging on target volume delineation in SBRT patients with central versus peripheral lung tumors. Multi-reader comparative study. Radiother Oncol. 2015;115:335-41.

[6] Kepka L, Socha J. PET-CT use and the occurrence of elective nodal failure in involved field radiotherapy for non-small cell lung cancer: A systematic review. Radiother Oncol. 2015;115:151-6.

[7] Aerts HJWL, Bussink J, Oyen WJG et al. Identification of residual metabolic-active areas within NSCLC tumours using a pre-radiotherapy FDG-PET-CT scan: A prospective validation. Lung Cancer. 2012;75:73-6.

[8] Calais J, Thureau S, Dubray B et al. Areas of High 18F-FDG Uptake on Preradiotherapy PET/CT Identify Preferential Sites of Local Relapse After Chemoradiotherapy for Non–Small Cell Lung Cancer. J Nucl Med. 2015;56:196-203.

[9] Bentzen SM. Theragnostic imaging for radiation oncology: dose-painting by numbers. Lancet Oncol. 2005;6:112-7.

[10] Usmanij EA, de Geus-Oei L-F, Troost EGC et al. 18F-FDG PET Early Response Evaluation of Locally Advanced Non–Small Cell Lung Cancer Treated with Concomitant Chemoradiotherapy. J Nucl Med. 2013;54:1528-34.

[11] Im H-J, Pak K, Cheon G et al. Prognostic value of volumetric parameters of 18F-FDG PET in non-small-cell lung cancer: a meta-analysis. Eur J Nucl Med Mol Imaging. 2015;42:241-51.

[12] Yossi S, Krhili S, Muratet J-P, Septans A-L, Campion L, Denis F. Early Assessment of Metabolic Response by 18F-FDG PET During Concomitant Radiochemotherapy of Non–Small Cell Lung Carcinoma Is Associated With Survival: A Retrospective Single-Center Study. Clin Nucl Med. 2015;40:e215-e21.

[13] Hyun S, Ahn H, Kim H et al. Volume-based assessment by 18F-FDG PET/CT predicts survival in patients with stage III non-small-cell lung cancer. Eur J Nucl Med Mol Imaging. 2014;41:50-8.

[14] de Geus-Oei L-F, Oyen WJG. Predictive and prognostic value of FDG-PET. Cancer Imaging. 2008;8:70-80.

[15] Zaidi H, El Naqa I. PET-guided delineation of radiation therapy treatment volumes: a survey of image segmentation techniques. Eur J Nucl Med Mol Imaging. 2010;37:2165-87.

[16] Lee JA. Segmentation of positron emission tomography images: Some recommendations for target delineation in radiation oncology. Radiother Oncol. 2010;96:302-7.

[17] Visser EP, Boerman OC, Oyen WJG. SUV: From Silly Useless Value to Smart Uptake Value. J Nucl Med. 2010;51:173-5.

[18] Konert T, Vogel W, MacManus MP et al. PET/CT imaging for target volume delineation in curative intent radiotherapy of non-small cell lung cancer: IAEA consensus report 2014. Radiother Oncol. 2015;116:27-34.

[19] Boellaard R, Willemsen AT, Arends B, Visser EP. EARL procedure for assessing PET/CT system specific patient FDG activity preparations for quantitative FDG PET/CT studies. Accessed Oct 2015.

[20] van Dalen JA, Hoffmann AL, Dicken V et al. A novel iterative method for lesion delineation and volumetric quantification with FDG PET. Nucl Med Commun. 2007;28:485-93.

[21] Daisne J-F, Sibomana M, Bol A, Doumont T, Lonneux M, Grégoire V. Tri-dimensional automatic segmentation of PET volumes based on measured source-to-background ratios: influence of reconstruction algorithms. Radiother Oncol. 2003;69:247-50.

[22] Hatt M, Cheze le Rest C, Turzo A, Roux C, Visvikis D. A Fuzzy Locally Adaptive Bayesian Segmentation Approach for Volume Determination in PET. IEEE Trans Med Imaging. 2009;28:881-93.

[23] Arens AIJ, Troost EGC, Hoeben BAW et al. Semiautomatic methods for segmentation of the proliferative tumour volume on sequential FLT PET/CT images in head and neck carcinomas and their relation to clinical outcome. Eur J Nucl Med Mol Imaging. 2014;41:915-24.

[24] van Elmpt W, De Ruysscher D, van der Salm A et al. The PET-boost randomised phase II dose-escalation trial in non-small cell lung cancer. Radiother Oncol. 2012;104:67-71.

[25] Ingerid Skjei K, Svestad JrG, Erlend Peter Skaug S et al. Validation of dose painting of lung tumours using alanine/EPR dosimetry. Phys Med Biol. 2016;61:2243.

[26] Even AJG, van der Stoep J, Zegers CML et al. PET-based dose painting in non-small cell lung cancer: Comparing uniform dose escalation with boosting hypoxic and metabolically active sub-volumes. Radiother Oncol. 2015;116:281-6.

[27] Knudtsen IS, van Elmpt W, Öllers M, Malinen E. Impact of PET reconstruction algorithm and threshold on dose painting of non-small cell lung cancer. Radiother Oncol. 2014;113:210-4.

[28] Gao A, Wang S, Fu Z, Sun X, Yu J, Meng X. (18)F-FDG avid volumes on pre-radiotherapy FDG PET as boost target delineation in non-small cell lung cancer. Int J Clin Exp Med. 2015;8:7561-8.

[29] Nygård L, Vogelius IR, Fischer BM et al. Early lesion-specific ^{18}F-FDG PET response to chemotherapy predicts time to lesion progression in locally advanced non-small cell lung cancer. Radiother Oncol. 2016;118:460-4.

[30] van Diessen JNA, Chen C, van den Heuvel MM, Belderbos JSA, Sonke J-J. Differential analysis of local and regional failure in locally advanced non-small cell lung cancer patients treated with concurrent chemoradiotherapy. Radiother Oncol. 2016;118:447-52.

Figure legends

Figure 1| Box and whisker plots of summed total lesion glycolysis (TLG) of the primary tumor and the lymph nodes in pre-treatment and interim ^{18}F-fluorodeoxyglucose (FDG) positron emission tomography (PET) scans obtained with different segmentation algorithms. Bottom and top of each box are lower and upper quartiles. The horizontal line near middle of the box is median. Whiskers are drawn down to the 5% percentile up to the 95% percentile, while the outliers are indicated by a dot. T40= Fixed level threshold segmentation at 40% of the maximum standardized uptake value (SUV_{max}), T50= Fixed level threshold segmentation at 50% of SUV_{max}, RTL= Relative level thresholding, SBR= Signal-to-background segmentation, FLAB=Fuzzy locally adaptive Bayesian segmentation.

Figure 2|Baseline (a+c) and early response monitoring (ERM) (b+d) ^{18}F-fluorodeoxyglucose (FDG) positron emission tomography (PET) images of two non-small cell lung cancer (NSCLC) patients. The first patient (a+b) showed a good response to treatment on both the primary tumour and lymph nodes. Although the primary tumour of the second patient (c+d) showed a good response to treatment, there was a limited response considering the lymph nodes, with more positive lymph nodes in the ERM PET. The mean summed fractional decrease of total lesion glycolysis (ΔTLG) of the first patient for the different segmentation methods was 76±6%, with a progression free survival (PFS) of 11 months and overall survival (OS) of 24 months. For the second patient, mean ΔTLG was 38±6% with a PFS and OS of 7 and 21 months, respectively.

Table

Tables and legends

Table 1| Patients Characteristics

Characteristics of patient population	
Male(Female)	18(9)
Median age (range) [y]	58 (42-77)
Histological type	
Non-small cell lung cancer (NSCLC)	60
Squamous cell carcinoma	10
Adenocarcinoma	14
NSCLC not otherwise specified	3
Disease stage[c]	
IIIA	520
IIIB	7
Performance-score (ECOG)	
0	19
1	7
Smoking status	
Current smoker	11
Former smoker	16
Lesion location	
Right upper lobe	11
Right middle lobe	4
Right lower lobe	2
Left upper lobe	8
Left lower lobe	2
Pre-treatment PET acquisition	
Number of bed positions	7-8
Administered FDG activity [MBq]	267±48
Incubation time [min]	75±7.5
Acquisition time per bed position [min]	4
Interim PET acquisition	
Number of bed positions	4-5
Administered FDG activity	269±49
Incubation time	78±8.0
Acquisition time per bed position [min]	4

Data are reported as mean±standard deviation. PET=positron emission tomography, FDG=^{18}F-fluorodeoxyglycose

Table 2| Hazard ratios (HRs) and 95% confidence interval (CI) of pre-treatment primary tumour total lesion glycolysis (TLG_T) and relative TLG decrease (ΔTLG_T) between pre-treatment and interim ^{18}F-fluorodeoxyglucose (FDG) positron emission tomography (PET) in a univariate Cox regression analysis for progression-free survival (PFS) and overall-survival (OS).

	HR (95% CI) per unit change for PFS	**Significance level**
Pre-treatment TLG_T		
T40	1.002 (1.000 – 1.004)	0.02*
T50	1.002 (1.000 – 1.004)	0.03*
RTL	1.002 (1.000 – 1.004)	0.03*
SBR	1.002 (1.000 – 1.003)	0.02*
FLAB	1.002 (1.000 – 1.003)	0.03*
ΔTLG_T		
T40	1.02 (1.003 – 1.03)	0.03*
T50	1.02 (1.003 – 1.03)	0.03*
RTL	1.02 (1.003 – 1.04)	0.03*
SBR	1.02 (1.004 – 1.04)	0.02*
FLAB	1.02 (1.000 – 1.04)	0.07
	HR (95% CI) per unit change for OS	**Significance level**
Pre-treatment TLG_T		
T40	1.002 (1.001 – 1.004)	0.004*
T50	1.003 (1.001 – 1.004)	0.005*
RTL	1.002 (1.001 – 1.004)	0.004*
SBR	1.002 (1.001 – 1.003)	0.004*
FLAB	1.002 (1.001 – 1.003)	0.006*
ΔTLG_T		
T40	1.02 (1.00 – 1.03)	0.05
T50	1.02 (1.00 – 1.03)	0.02*
RTL	1.02 (1.00 – 1.03)	0.08
SBR	1.02 (1.00 – 1.03)	0.04*
FLAB	1.01 (0.99 – 1.03)	0.3

T40= fixed level threshold at 40% of the maximum standardized uptake voxel (SUV_{max}), T50= fixed level threshold at 50% of SUV_{max}, RTL=relative level thresholding, SBR=signal-to-background-ratio, FLAB=fuzzy locally adaptive Bayesian segmentation. Statistical significance is indicated by an asterisk '*'.

Table 3| Hazard ratios (HRs) and 95% confidence interval (CI) of pre-treatment summed total lesion glycolysis (TLG_S) and relative TLG decrease (ΔTLG_S) between pre-treatment and interim ^{18}F-fluorodeoxyglucose (FDG) positron emission tomography (PET) in a univariate Cox regression analysis for progression-free survival (PFS) and overall-survival (OS). TLG_S is the sum of primary tumour TLG (TLG_T) and lymph node TLG (TLG_{LN}).

	HR (95% CI) per unit change for PFS	**Significance level**
Pre-treatment TLG_S		
T40	1.003 (1.001 – 1.005)	0.002*
T50	1.003 (1.001 – 1.005)	0.004*
RTL	1.003 (1.001 – 1.005)	0.003*
SBR	1.002 (1.001 – 1.004)	0.004*
FLAB	1.002 (1.001 – 1.004)	0.004*
ΔTLG		
T40	1.02 (1.00 – 1.05)	0.03*
T50	1.03 (1.02 – 1.05)	0.001*
RTL	1.01 (1.00 – 1.02)	0.2
SBR	1.04 (1.02 – 1.06)	0.001*
FLAB	1.02 (1.00 – 1.04)	0.04*
	HR (95% CI) per unit change for OS	**Significance level**
Pre-treatment TLG_S		
T40	1.003 (1.001 – 1.004)	0.001*
T50	1.003 (1.001 – 1.005)	0.002*
RTL	1.003 (1.001 – 1.005)	0.001*
SBR	1.002 (1.001 – 1.004)	0.001*
FLAB	1.002 (1.001 – 1.004)	0.002*
ΔTLG_S		
T40	1.01 (1.00 – 1.03)	0.4
T50	1.02 (1.00 – 1.04)	0.02*
RTL	1.00 (1.00 – 1.01)	0.9
SBR	1.02 (1.00 – 1.04)	0.03*
FLAB	1.01 (1.00 – 1.03)	0.02*

T40= fixed level threshold at 40% of the maximum standardized uptake voxel (SUV_{max}), T50= fixed level threshold at 50% of SUV_{max}, RTL=relative level thresholding, SBR=signal-to-background-ratio, FLAB=fuzzy locally adaptive Bayesian segmentation. Statistical significance is indicated by an asterisk '*'.

Figure1

Click here to download high resolution image

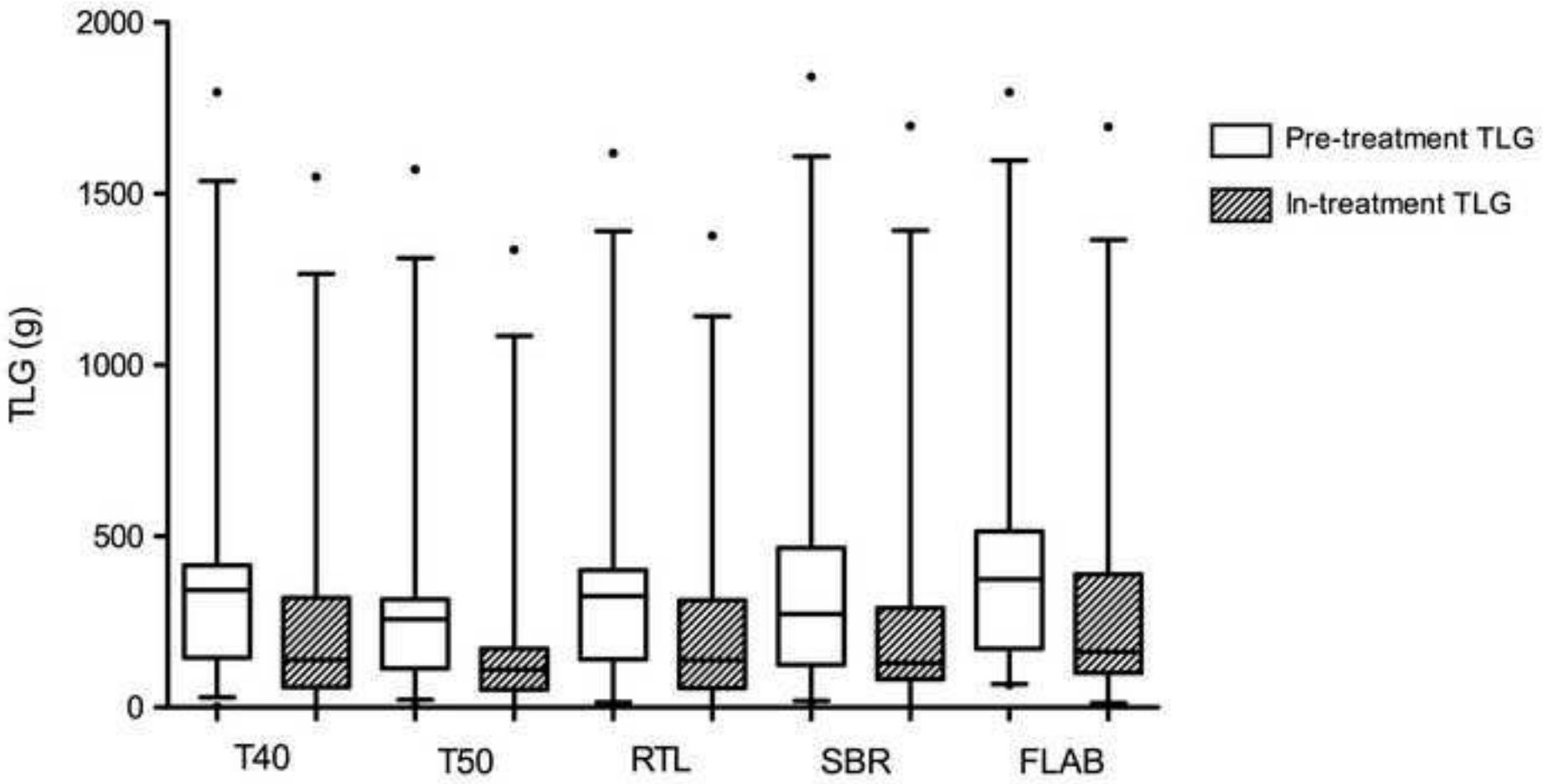


Figure2a

Click here to download high resolution image

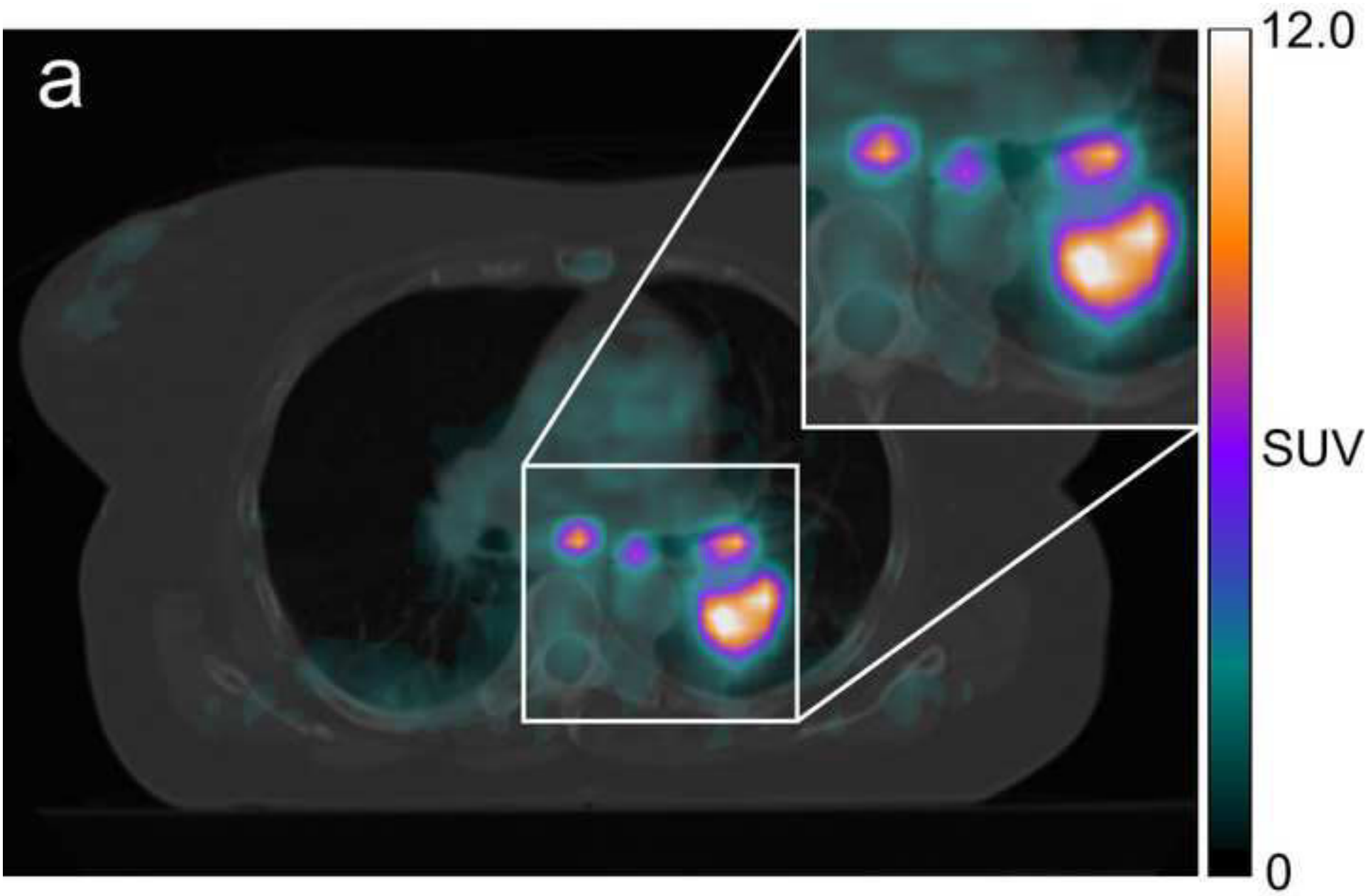


Figure2b
Click here to download high resolution image

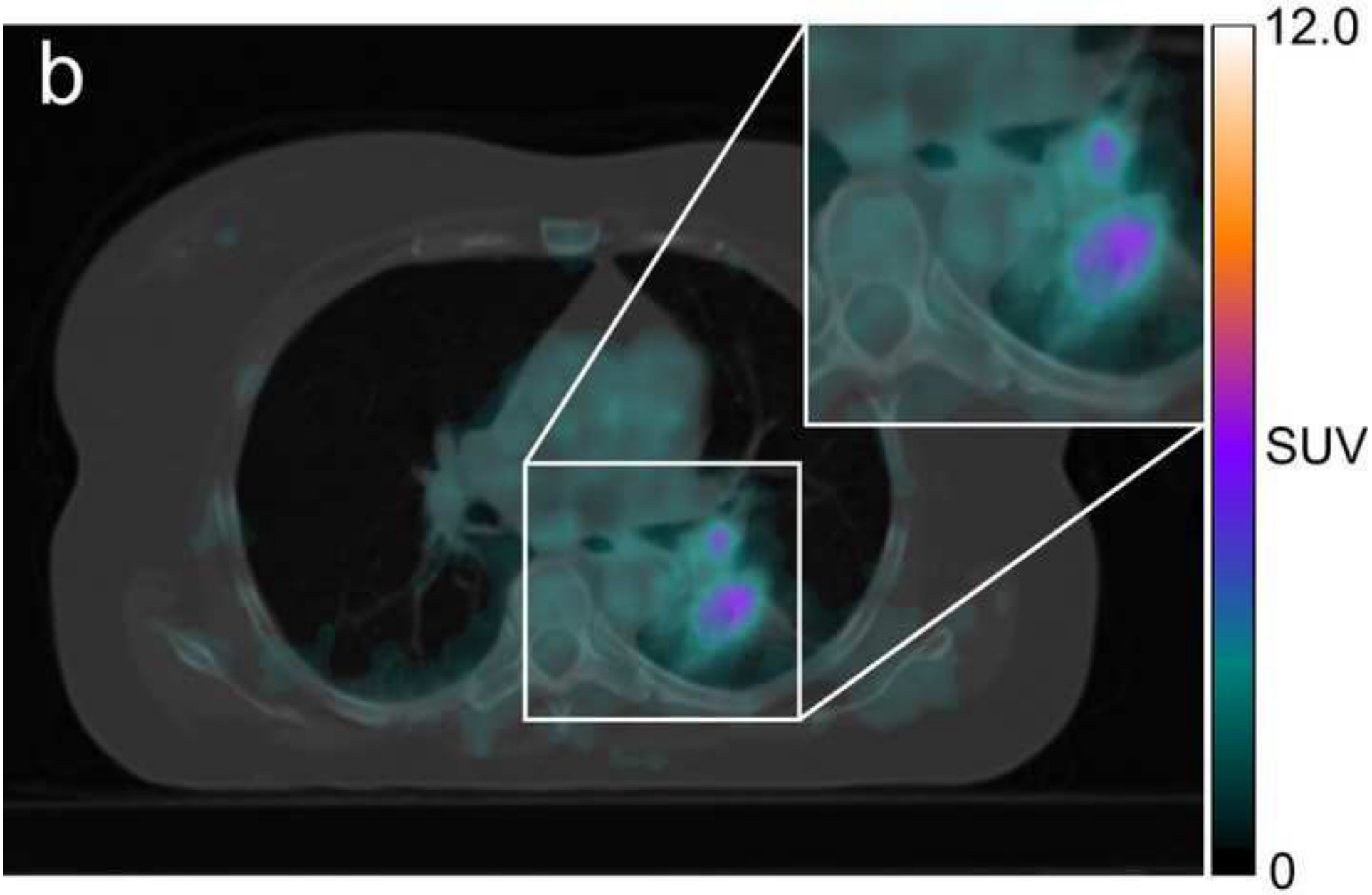


Figure2c
Click here to download high resolution image

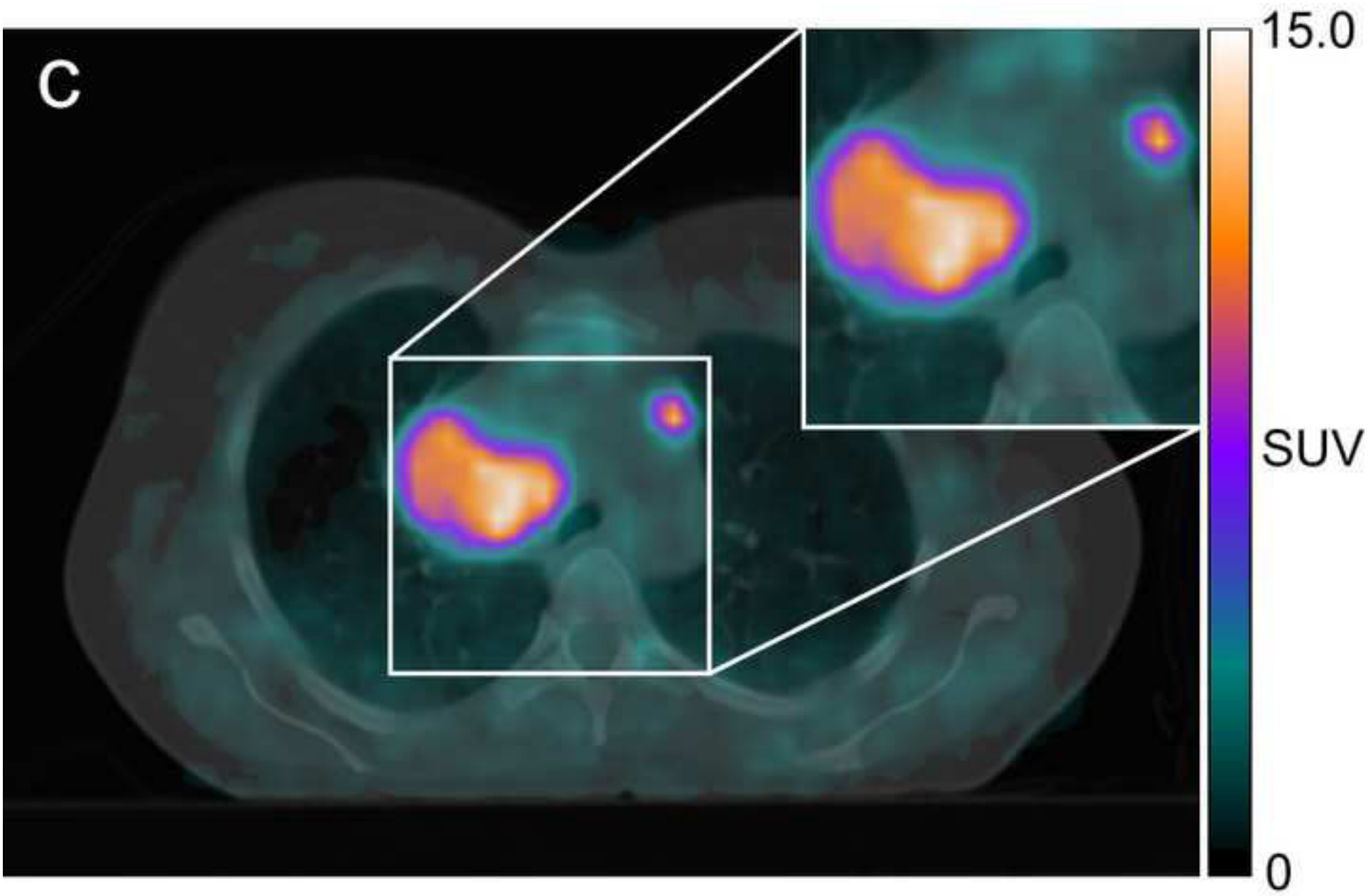
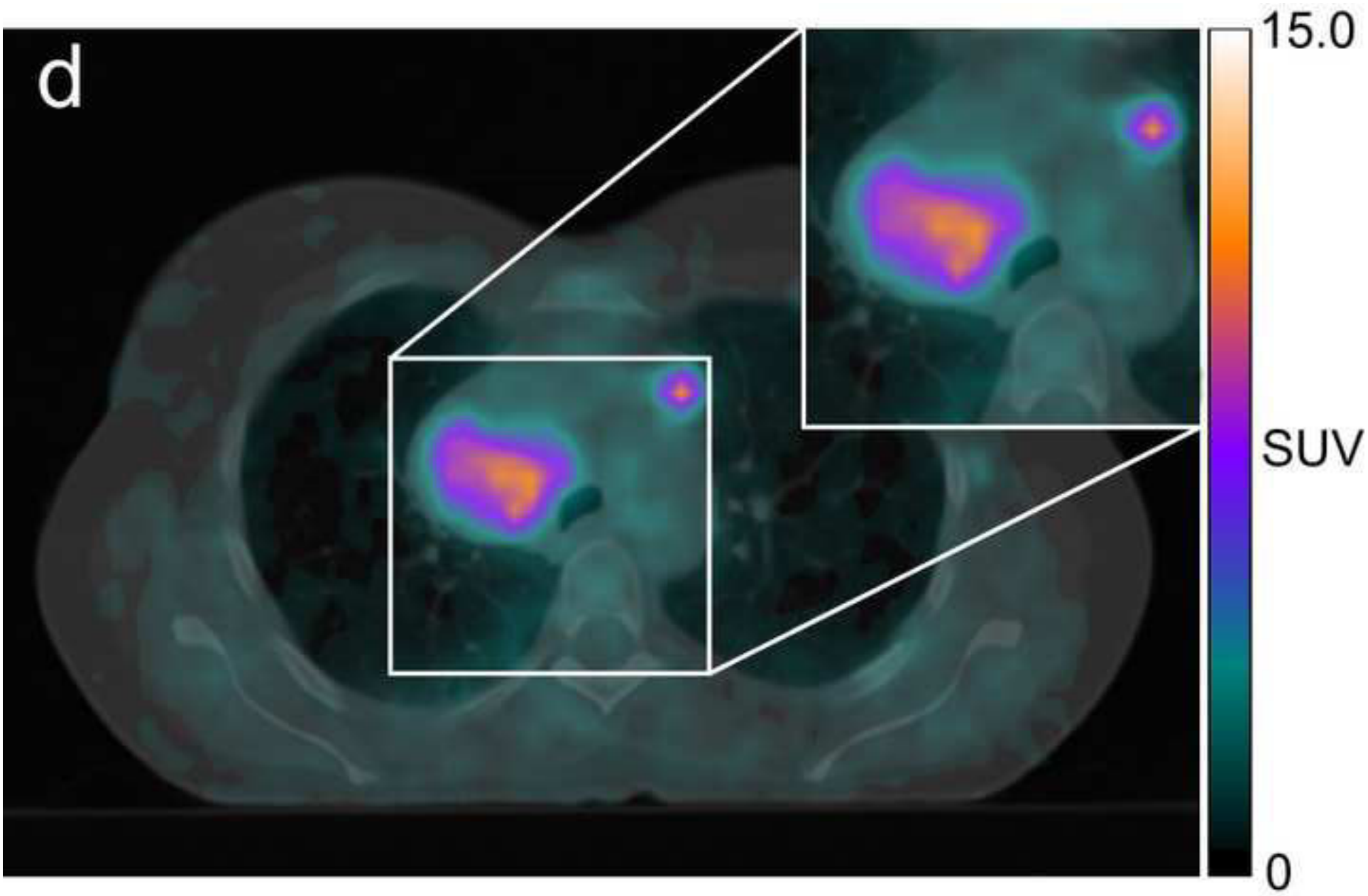
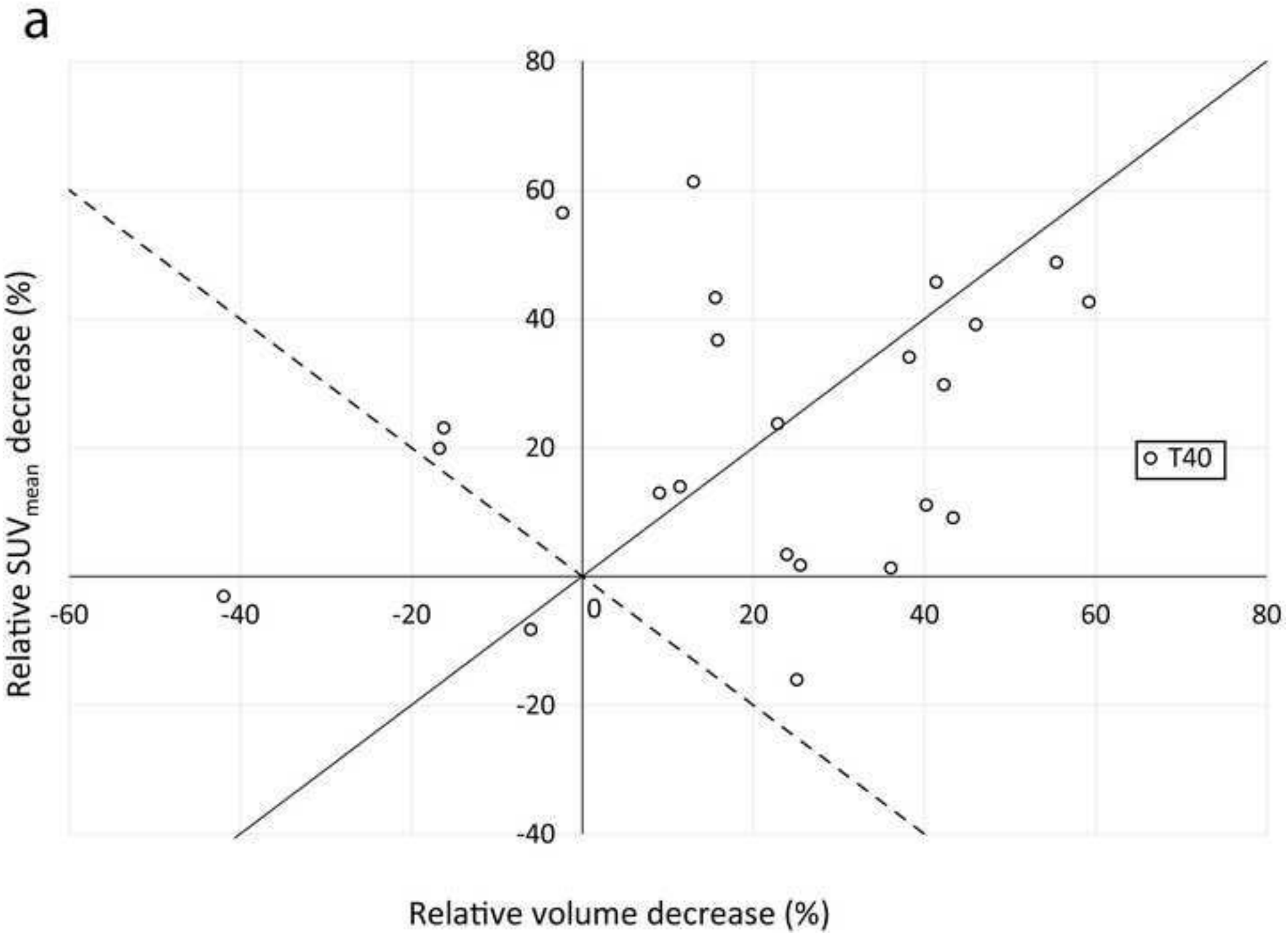


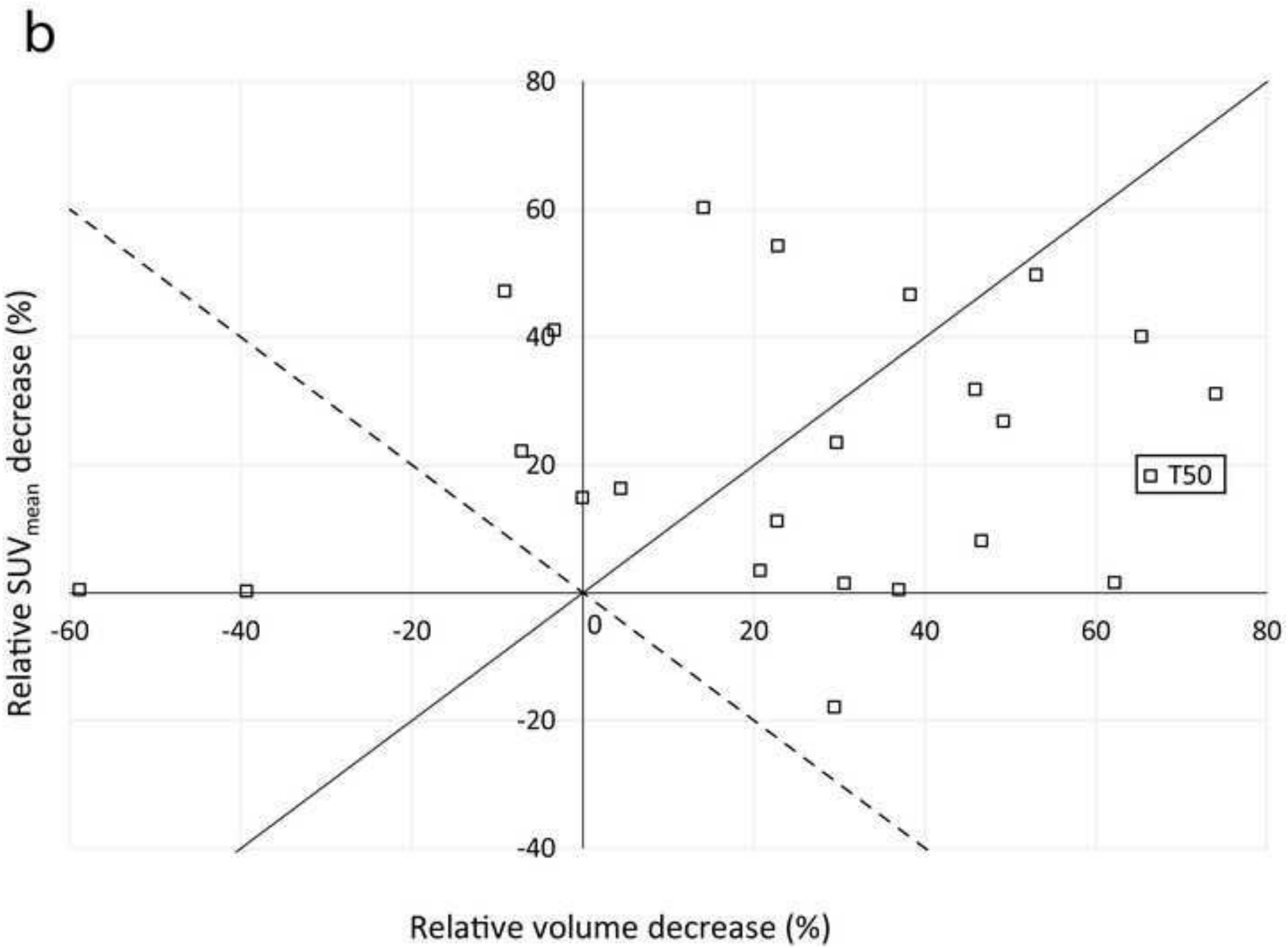
Figure2d
Click here to download high resolution image



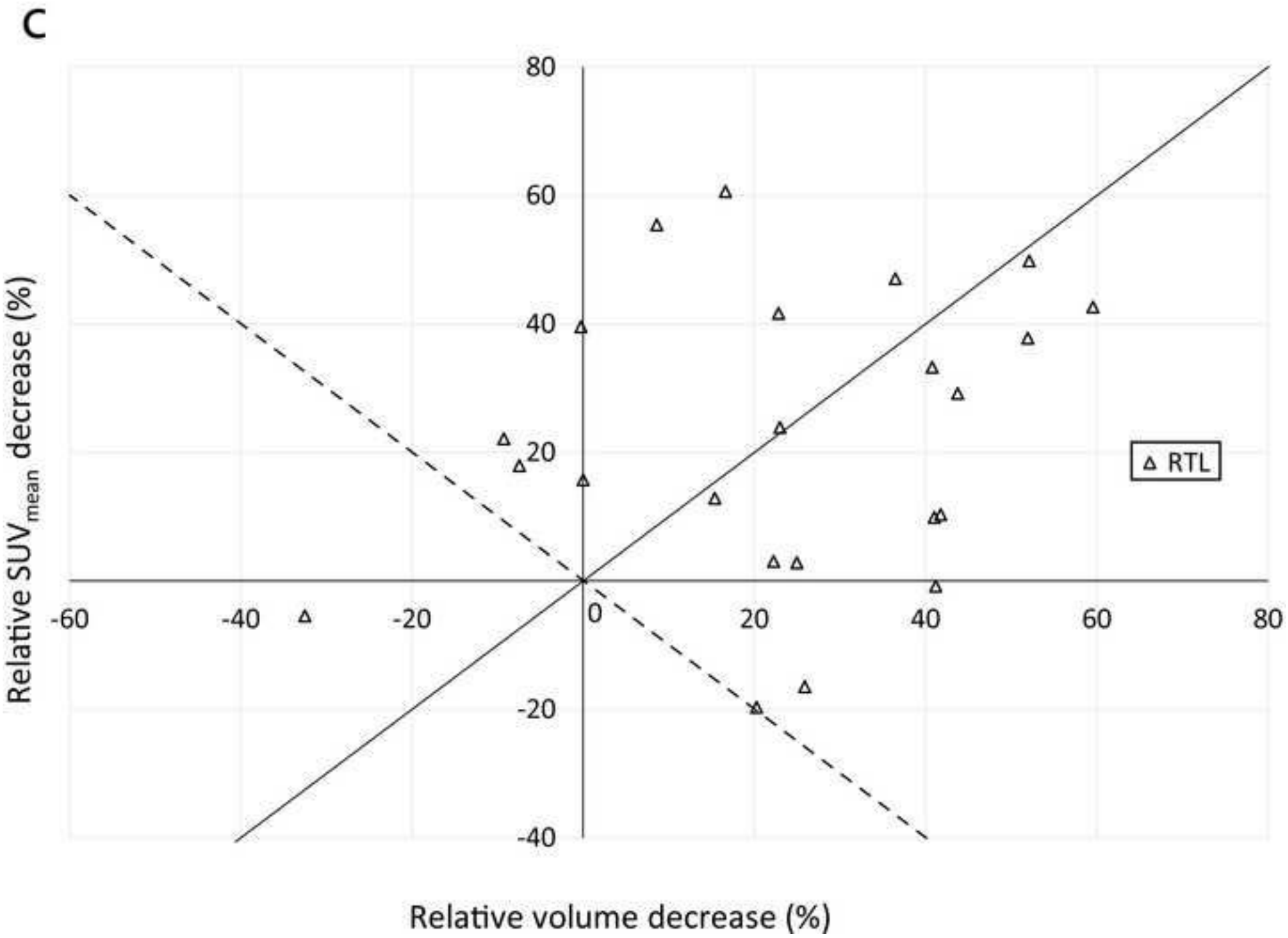
Supplementary figure 1a
Click here to download high resolution image



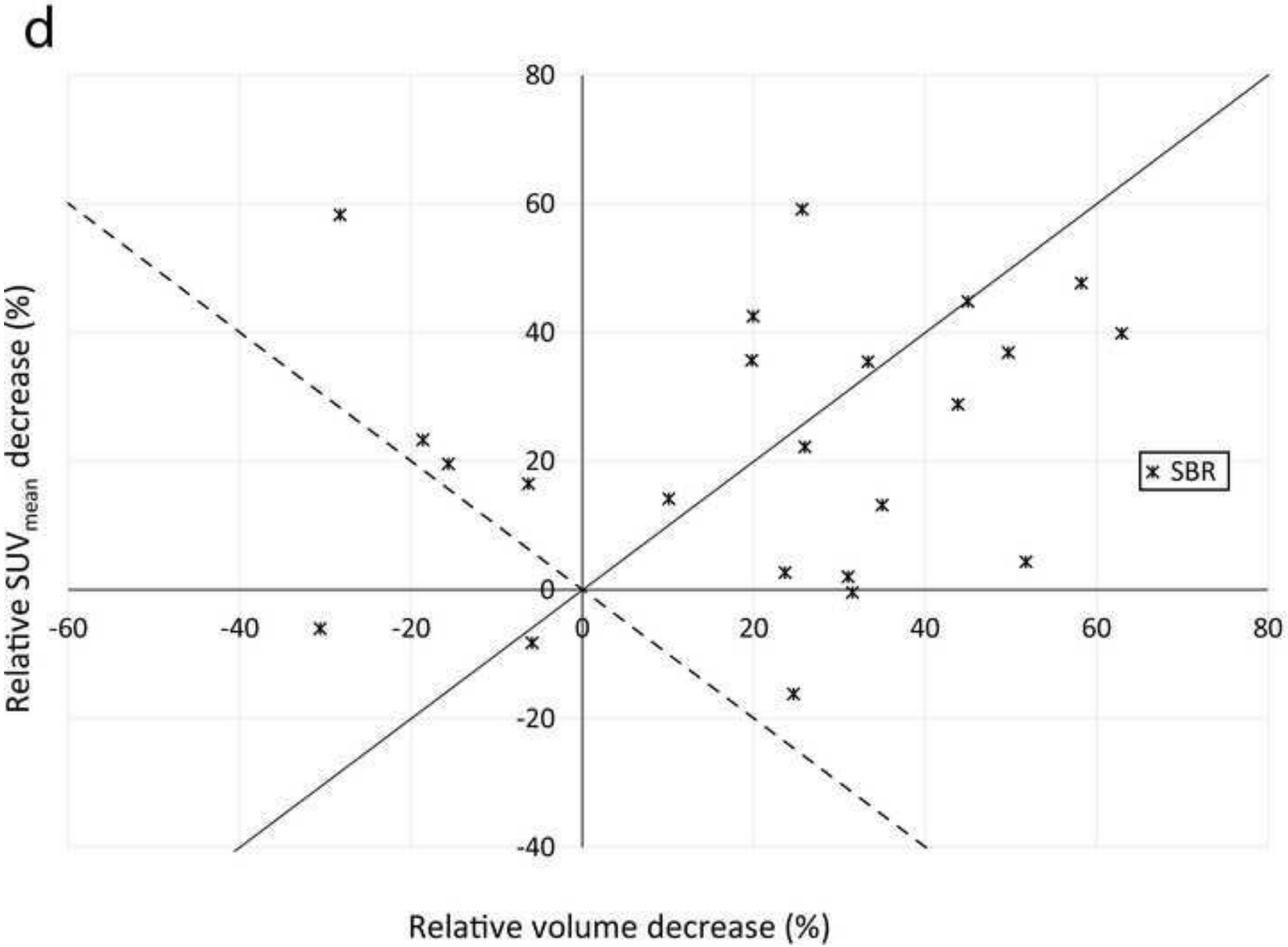
Supplementary figure 1b
Click here to download high resolution image



Supplementary figure 1c
Click here to download high resolution image



Supplementary figure 1d
Click here to download high resolution image



Supplementary figure 1e
Click here to download high resolution image