View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by HAL-Université de Bretagne Occidentale

Impact of partial-volume effect correction on the predictive and prognostic value of baseline 18F-FDG PET images in esophageal cancer.

Mathieu Hatt, Adrien Le Pogam, Dimitris Visvikis, Olivier Pradier, Catherine Cheze Le Rest

▶ **To cite this version:**

Mathieu Hatt, Adrien Le Pogam, Dimitris Visvikis, Olivier Pradier, Catherine Cheze Le Rest. Impact of partial-volume effect correction on the predictive and prognostic value of baseline 18F-FDG PET images in esophageal cancer.: Partial volume correction on FDG PET. Journal of Nuclear Medicine, Society of Nuclear Medicine, 2012, 53 (1), pp.12-20. <10.2967/jnumed.111.092775>. <inserm-00707272>

HAL Id: inserm-00707272

http://www.hal.inserm.fr/inserm-00707272

Submitted on 12 Jun 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Impact of partial volume effects correction on the predictive and prognostic value of baseline ^{18}F-FDG PET images in esophageal cancer

Mathieu Hatt[1], Adrien Le Pogam[1,2], Dimitris Visvikis[1], Olivier Pradier[1,3], Catherine Cheze Le Rest[1]

[1] INSERM, UMR 1101 LaTIM, CHRU Morvan, Brest, France

[2] MRC Clinical Sciences Centre, Hammersmith Hospital, London, UK

[3] Department of radiotherapy, CHRU Morvan, Brest, France

Running title: Partial volume correction on FDG PET

Corresponding author:

Mathieu HATT,
LaTIM, INSERM UMR 1101, CHRU MORVAN,
5 avenue Foch, 29609 Brest,
France
Tel.:+33298018111
Fax:+33298018124:

hatt@univ-brest.fr

Wordcount: 5636

Keywords: esophageal cancer, response to therapy, overall survival, PET, partial volume effects, SUV, tumor volume, total lesion glycolysis

ABSTRACT

The objectives of this study were to investigate the clinical impact of partial volume effects (PVE) correction on the predictive and prognostic value of metabolically active tumor volume (MATV) measurements on 18F-FDG PET baseline scan for therapy response and overall survival in esophageal cancer patients.

Methods: 50 patients with esophageal cancer treated with concomitant radio-chemotherapy between 2004 and 2008 were retrospectively considered. PET baseline scans were corrected for PVE with iterative deconvolution incorporating wavelet denoising. MATV delineation on both original and corrected images was carried out using the automatic Fuzzy Locally Adaptive Bayesian (FLAB) methodology. Several parameters were extracted considering the original and corrected images: maximum and peak SUV, mean SUV, MATV and TLG (TLG=MATV×mean SUV). The predictive value of each parameter with or without correction was investigated using Kruskal-Wallis tests and the prognostic value with Kaplan-Meier curves.

Results: Whereas PVE correction had significant quantitative impact on the absolute values of the investigated parameters, their clinical value within the clinical context of interest was not significantly modified. This was observed for both overall survival and response to therapy. The hierarchy between parameters was the same before and after correction. SUV measurements (max, peak, mean) had non-significant (p>0.05) predictive or prognostic value, whereas functional tumor related measurements (MATV, TLG) were significant (p<0.002) predictors of response and independent prognostic factors.

Conclusions: PVE correction does not improve the predictive and prognostic value of baseline PET image derived parameters in esophageal cancer patients.

With a worldwide estimated 5-year survival of only 15% (*1*), esophageal cancer is the third most common malignancy of the digestive tract and is a leading cause of cancer mortality. Its incidence is still increasing and there is a growing concern regarding its effective management (*2*). Surgical resection remains the most effective treatment, however many patients have a locally advanced esophageal carcinoma (LAEC) at diagnosis and neoadjuvant therapy before surgery has demonstrated improved survival in this case (*3*). The maximum improvement in terms of increased overall survival from neoadjuvant treatment is observed for patients who achieve a complete pathological response (only 15-30% of cases) with no residual cancer cells in the primary tumor or lymph nodes (*4*). On the other hand, non responders may be unnecessarily affected by toxicity (*5*). The development of an early diagnostic test offering non invasive prediction of the response to therapy and/or survival is therefore of great interest. For tumors that cannot be surgically removed, combined radio-chemotherapy is the preferred treatment. In this case too, early assessment of response to therapy would allow a modification in the management of non responding patients early during treatment. Such a response assessment becomes even more critical when one considers the availability of new targeted drugs that could be tested with higher efficiency if applied early (*6*).

Along with Standardized Uptake Values (SUV, max or peak) usually considered in clinical practice, other parameters describing functional lesions, such as metabolically active tumor volume (MATV, defined as the tumor volume that can be seen and delineated on an ^{18}F-FDG PET image) (*7*), mean SUV and total lesion glycolysis (TLG, defined as the product of MATV and its associated SUV_{mean}) (*8*) have been investigated. The prognostic value of these parameters in esophageal cancer patients for overall or disease-free survival has been demonstrated (*9-12*). On the

other hand regarding therapy prediction, several studies on different cancer models have recently suggested using the baseline scan only, instead of the comparison of pre-treatment and post-treatment scans (late assessment) or during-treatment scans (early assessment) (*13*). Such investigations were for instance carried out in pleural mesothelioma (*14*), non-Hodgkin lymphoma (*15*) and esophageal cancer (*7, 16*), demonstrating higher statistical value for MATV-based parameters than SUV measurements whose predictive value has been found to be conflicting (*17*).

However, in most of these studies no partial volume effects (PVE) correction (PVC) was applied which may explain the observed limited value of SUV. The impact of PVC on the clinical value of SUV measurements has been investigated by a limited number of authors. Hoetjes et al (*18*) investigated the impact of four PVC strategies on 15 breast cancer patients, regarding the early metabolic PET response after one cycle of chemotherapy. The SUV decrease between the pre-treatment scan and the scan early during treatment was found to be lower after PVC (26-27% vs. 31%) for the first three methods but not for the fourth one based on binary tumor masks (30%). Van Heijl et al (*19*) recently demonstrated a non-significant impact of PVC on the correlation between disease-free survival and ^{18}F-FDG PET SUV measurements in 52 esophageal cancer patients. In this study a PVC method based on binary tumor masks generated using adaptive thresholding delineation was used, while disease-free survival was the only clinical endpoint investigated. Both the use of adaptive thresholding and the PVC method based on tumor masks assume a homogeneous tracer distribution in both tumor and background and are therefore likely to provide only approximate correction (*20*). On the other hand, no data is currently available regarding the impact of PVC on the value of baseline ^{18}F-FDG PET based

measurements for the prediction of overall survival and response to therapy in esophageal cancer.

The current study was therefore carried out to investigate the impact of an advanced PVC methodology and the use of an accurate MATV delineation approach on both the predictive and prognostic value of baseline ^{18}F-FDG PET scan derived parameters.

MATERIAL AND METHODS

Patients

50 consecutive patients with a newly diagnosed esophageal cancer were included and retrospectively analyzed. The characteristics of the patients are given in table 1. Most of them (45 out of 50) were male, aged 65±9 years at the time of diagnosis. 74% of the tumors originated from the middle and lower esophagus and 72% were squamous cell carcinoma. None of the patients underwent surgery, and all were treated with concomitant radio-chemotherapy between 2004 and 2009. The therapy regime included three courses of 5-fluorouracil/cisplatin and a median radiation dose of 60Gy given in 180cGy daily fractions delivered once daily, 5 days a week for 6-7 weeks. As part of the routine procedure for the initial staging in esophageal cancer, each patient was referred for an ^{18}F-FDG PET study before treatment, and these baseline scans were used in this study.

Overall survival was determined as the time between initial diagnosis and last follow-up or death. Response to therapy was evaluated one month after the completion of the concomitant radio-chemotherapy using conventional thoraco abdominal CT and endoscopy. Patients were classified as non responders (NR, including stable and progressive disease), partial responders (PR) or complete responders (CR).

Response evaluation was based on CT evolution between pre-treatment and post-treatment scans using RECIST (Response Evaluation Criteria in Solid Tumors) (*21*). Patients also underwent fibroscopy in case of partial or complete response. Complete response was confirmed by the absence of visible disease in the endoscopy and no viable tumor on biopsy. Partial CT response was confirmed by macroscopic residual (>10% viable) on biopsy. No discordance was observed between pathological, when available, and CT evaluation. The current analysis was carried out after an approval by the institutional ethics review board.

^{18}F-FDG PET acquisitions

^{18}F-FDG PET studies were carried out prior to the treatment. Patients were instructed to fast for at least 6h before a 5MBq/kg injection of ^{18}F-FDG. Static emission images were acquired from head to thigh beginning 60min after injection and with 2min per bed position, on a Philips GEMINI PET/CT system (Philips Medical Systems, Cleveland, OH USA). Images were reconstructed using the RAMLA 3D algorithm according to standard clinical protocol: 2 iterations, relaxation parameter of 0.05, a 5mm 3D Gaussian post-filtering, a 4x4x4mm^3 voxels grid sampling, and a low dose CT scan-based attenuation correction.

PET image partial volume correction and image analysis

Images were corrected for PVE using an iterative deconvolution methodology that has been previously validated (*22*). Its principle is to iteratively estimate the inversion of the scanner's Point Spread Function (PSF), which is assumed to be known and spatially invariant in the field of view. The considered lesions were all in the same body region and this approximation should therefore not have a significant impact on the applied correction on a patient-by-patient comparison basis. Iterative deconvolution methods, such as Lucy-Richardson (L-R) or Van Cittert, are known for

the amplification of noise associated with increasing number of iterations. To solve this issue, wavelet-based denoising of the residual was introduced within the iterative L-R deconvolution using Bayeshrink filtering (*23*), leading to images corrected for PVE without significant noise addition. The advantages of this methodology are its ability to generate entire whole-body corrected images independently of any manual or automatic segmentation of regions of interest. It is also voxel-based and therefore does not assume homogeneous regional radiotracer distributions for the tumor and/or surrounding background.

Tumor delineation and parameters extraction

For each patient, the tumor was identified on the baseline pretreatment PET images by an experienced nuclear physician. It was then delineated using the Fuzzy Locally Adaptive Bayesian (FLAB) algorithm (*20, 24*) on both the original (without PVE correction) and PVE corrected images. This segmentation approach has been shown to give both robust and reproducible functional volume delineations under variable image noise characteristics (*25-26*).

The following parameters were subsequently extracted from each baseline image with or without correction for PVE: SUV_{max}, SUV_{peak} defined as the mean of SUV_{max} and its 26 neighbors (roughly corresponding to a 1cm ROI), mean SUV (SUV_{mean}) within the volume, MATV, and TLG (determined by multiplying SUV_{mean} with the corresponding MATV).

Statistical analysis

Pearson coefficients were used to estimate correlation between the image derived parameters, and paired t-tests were used to characterize the differences between uncorrected and corrected parameters. The correlation between response to therapy and each parameter was investigated using the Kruskal-Wallis test as a non-

parametric statistic allowing the comparison of parameter distributions associated to each category of response (CR, PR and NR). This test does not assume normal distribution of variables and the computation of its statistic H is based on ranks instead of absolute values of variables (*27*). Regarding survival, for each considered parameter, Kaplan-Meier survival curves were generated (*28*) for which the most discriminating threshold value allowing differentiation of the groups of patients was identified using receiver operating characteristic (ROC) methodology (*29*). Prognostic value of each parameter in terms of overall survival was assessed by the log-rank test.

The significance of the following factors (with or without correction) was tested: SUV_{max}, SUV_{peak}, MATV, SUV_{mean}, and TLG. All tests were performed two-sided using the Medcalc^{TM} statistical software (MedCalc Software, Belgium) and *p* values below 0.05 were considered statistically significant.

RESULTS

Impact of PVC on the image derived parameters

The PVE correction had an impact on the images that can be assessed visually, with a higher contrast between the tumor and the surrounding tissues, as it can be seen in figure 1 and illustrated using profiles in figure 2. Table II provides the distributions of volumes and associated parameters measured in original and corrected images.

MATV delineated on original images and these delineated on images corrected for PVE were highly correlated (r>0.998, CI 0.997-0.999, $p<0.0001$). However, MATV delineated on PVC corrected images were systematically smaller ($p<0.001$) by on average -10±5% (range -1.5 to -22.4%), which resulted in a mean volume difference of -4±3 cm^3 (40±36 cm^3 vs. 36±34 cm^3). This is illustrated on two different tumors in

figure 3. There was no significant correlation between these differences and the PET lesion volumes ($r<0.2$, $p>0.18$).

All primary lesions were detected by ^{18}F-FDG PET and exhibited a rather high uptake with a mean SUV_{max} of 10±4. As expected, SUV_{peak} and SUV_{mean} measurements were comparatively lower (8±3 and 6±2 respectively). All SUV measurements are summarized in table II. After iterative deconvolution, SUV_{max}, SUV_{peak} and SUV_{mean} were 15±6, 10±4 and 7±3 respectively. All were significantly higher than non-corrected values ($p<0.05$). SUV_{max} increased by 54±23% (range 18-157%) while the impact on SUV_{peak} and SUV_{mean} was lower with a mean increase of 27±10% (range 8-51%) and 28±11% (range 9-59%) respectively. Considering the PVC induced decrease of MATV (-10±5%) and increase of corresponding SUV_{mean} (+28±11%), PVC resulted in significantly higher TLG values (+14±12%, range -2 to +50%) ($p<0.0001$).

The increases of SUV_{max} and SUV_{peak} after PVC were not correlated with MATV ($r<0.2$, $p>0.2$), whereas the increase of SUV_{mean} was inversely correlated with MATV ($r=-0.79$, $p<0.0001$), with higher increase observed for smaller volumes.

Impact of PVC on the predictive and prognostic values

25 patients were classified as PR, 11 were CR and 14 were NR (including stable and progressive disease). With a median follow-up of 60 months (range 10-84), the median overall survival was 12 months and the 1-year and 2-year survival rates were 60% and 35% respectively. 10 patients were alive with no evidence of disease at the time of last follow-up, while 9 were alive with recurrent disease and 31 had died. Survival was significantly correlated with response, as overall survival was 24±15 (median 21), 22±20 (median 14) and 9±4 (median 10) months for CR, PR and NR

respectively (p<0.01). Results concerning the prognostic and predictive value of all considered parameters with and without PVC are summarized in tables III and IV.

Initial SUV_{max} whether corrected for PVE or not, was not predictive of response to therapy (*p*=0.2 and *p*=0.3 for SUV_{max} and SUV_{max}^{PVC} respectively) although CR tend to have smaller SUV_{max} (7.8±4.2 and 12.2±6.6 after PVC) than PR and NR (10.2±3.7 and 10.3±3.8 for PR and NR respectively, 15.9±6.0 and 15.5±5.7 after PVC) (figure 4A). SUV_{peak} led to slightly more differentiated groups of response without reaching statistical significance (*p*=0.08), with a mean value of 6.2±3.6 in complete responders, whereas both PR and NR were characterized by similar higher SUV_{peak} values (8.5±3.1 and 8.5±3.2 for PR and NR respectively). After PVC, the results using SUV_{peak} were similar with 7.8±4.4, 10.7±3.7 and 10.8±3.9 for CR, PR and NR respectively (*p*=0.1). The SUV_{mean} measurements could not significantly predict response (*p*=0.07), and the differentiation between the three groups of response considered based on SUV_{mean} was still not possible after PVC (*p*>0.14).

None of the SUV measurements was a significant prognostic factor in the univariate analysis, despite a trend for longer survival associated with lower SUV (max, peak or mean). For instance, a SUV_{max} below a threshold of 8 or a mean SUV under 6.5 tend to be associated with a better outcome and a median survival of 20 vs. 13 months (*p*=0.3) and 16 month vs. 10 months (*p*=0.15) respectively. Similarly after PVC no threshold value could significantly differentiate groups of patients regarding their survival (see figure 5A-B).

Contrary to SUV measurements with or without PVC, the parameters related to functional volume (MATV and TLG) allowed significant (*p*<0.0001) differentiation of the three response groups, and were significant prognostic factors (*p*<0.002), as

illustrated in figure 4C. No significant differences were found using the original or PVC corrected values.

The parameter that allowed for the best differentiation of patients groups was the TLG (p<0.0001). CR were characterized by a TLG of 55±45g, whereas PR and NR had a TLG of 178±143g and 416±238g respectively. After PVC, the absolute values of each group rose to 62±45g, 200±155g and 437±249g for CR, PR and NR respectively, leading to the same discrimination between groups of response (p<0.0001). Although slightly less efficient than TLG, the use of MATV allowed a statistically significant differentiation of the three response groups (p<0.0001). Using the MATV values extracted from PVC images led to exactly the same discriminating power (p<0.0001).

MATV and TLG were also good prognostic factors, with high MATV and TLG values being significantly associated with shorter survival with hazard ratios between 3 and 4 (table III). A MATV above 85cm^3 was identified as a predictor of poor outcome with a median survival of 6 months only vs. 20 months for patients with smaller MATV (p=0.0004) as illustrated in figure 5C. In addition, a MATV below 15cm^3 was associated (p=0.009) with longer survival (49 months) than larger MATV (11 months). Similar results were obtained using the MATVs measured on the PVC corrected images, with a median survival of 20 months for patients with TV^{PVC} below 80cm^3 vs. 10 months (p<0.002). Regarding TLG, a threshold of 260g was found to be a good discriminating factor for outcome (21 vs. 10 months, p=0.0012) whereas using PVC corrected TLG led to similar results with a slightly higher threshold (TLG^{PVC}=280g, 21 vs. 10 months, p=0.0004).

DISCUSSION

Our study investigated the impact of partial volume effects correction on the predictive and prognostic value of different parameters derived using the baseline pre-treatment PET images. Our results confirmed that PVE correction has a significant impact on quantitative SUV values with an average increase of above 50% for SUV_{max}, in agreement with previous studies (*18-19*), and lower increase (below 30%) for SUV_{peak} and SUV_{mean}. The lower increase observed for SUV_{peak} and SUV_{mean} is related to the fact that the L-R deconvolution is a voxel-by-voxel process leading to enhancement of contrasts between sub-volumes within the MATV and both lower and higher voxels' SUV values included in the averaging associated with the calculation of mean and peak SUV. PVC did not have a significant impact on the delineation of the MATV. Overall, MATVs delineated on the corrected images were only slightly smaller than those determined on the original images. The mean reduction of 10% was within the reproducibility limits of confidence intervals regarding tumor volume measurements on double baseline PET scans using FLAB (±30%) (*26*). This limited impact of PVC on MATV can be explained by the fact that PVE is dependent on tumor size and is more pronounced on small lesions (*30*). In our group of patients the tumors were rather large (40±30cm^3) therefore the relative variation of volumes with respect to the entire volume is small. 12 (25%) of them had MATV around 10cm^3 or smaller. In addition, the use of a robust delineation approach instead of threshold-based methods in various configurations of blur and noise (*25, 31*) ensured a limited variability in the MATV delineation results between original and corrected images.

As previously demonstrated (*7, 12*), MATV and TLG extracted from non corrected ^{18}F-FDG PET pre-treatment acquisitions had high clinical value. On the contrary,

none of the usual SUV measurements (max, peak or mean) considered in clinical practice was significantly associated with therapy response or survival, as also reported in the two largest available prospective trials (*32-33*).

Regarding response to therapy prediction using SUVs, we found that PVE correction did not improve the already demonstrated low discriminating power of any of the SUV measurements considered (*7*). This can be explained by the combination of several factors. First, without PVE correction, the trend of low SUV being associated with better outcome may have been exaggerated by an underestimation of SUV, since complete responders had also smaller volumes in addition to low SUV_{max}. Secondly, after PVE correction all three response groups had increased SUV_{max}, but with still no significant difference between the groups. We have demonstrated that mean SUV increase after PVC was inversely correlated with TV (r=0.8, p<0.0001), smaller volumes being characterized by higher mean SUV increases after PVC compared to larger volumes. The mean SUV values within the MATV of partial or non responders were therefore increased by a smaller amount (+20±9%) than those within the MATV of complete responders (+34±13%) which were associated with smaller tumor volumes. The mean tumor SUVs of complete responders were therefore closer to the mean SUVs of partial and non responders after correction. Hence, the discriminating power of SUV_{mean} was reduced by PVE correction. A similar trend was observed for SUV_{max} and SUV_{peak}, although it was less significant since their respective increase was not correlated to the MATV. Therefore PVC might have further reduced the clinical value of SUV measurements in this context. This effect has been previously suggested as a limitation to the prognostic value of SUV_{max} in early stage non small cell lung cancer (*34*).

Similar conclusions can be drawn from the results regarding the impact of PVC on the prognostic value of the SUV parameters. Indeed, as already demonstrated (*12*), extreme MATV values were significantly associated with longer or shorter overall survival for very small (49 months for MATV below 15 cm^3 vs. 11 months) or very large MATV (8 months for TV above 85cm^3 vs. 20 months) respectively. On the other hand, SUV measurements without correction cannot significantly differentiate between the patients with longer or shorter survival (p>0.05 for all SUV measurements) although a trend for longer survival was associated with lower SUVs. After correction, this differentiation was not significantly improved, because SUVs associated with the smaller volumes were closer to the SUVs associated with larger volumes. Therefore the discrimination was again reduced by PVC. To our knowledge there is no similar data available on the impact of PVC on SUV predictive value in the literature, but our results are in agreement with previous findings that demonstrated no significant changes in disease-free survival correlation between original and corrected SUVs in esophageal cancer using alternative less accurate methodologies for both PVE correction and functional volume segmentation (*19*).

As previously demonstrated (*7, 12*), MATV and associated TLG values were good predictors of response (*7*) and independent prognostic factors of overall survival (*12*). After PVC the already high clinical value of MATV and TLG was not significantly altered. Considering the thresholds used to differentiate patient groups, there was no need for adjustment regarding MATV measurements since MATVs were not significantly modified by PVC. On the other hand, TLG thresholds needed to be adjusted, considering that PVC led to significantly increased mean SUVs and resulting TLG values. The determined threshold values for each parameter regarding prognosis or prediction of response were found using ROC analysis on the current

patient cohort and would therefore require larger prospective studies in order to be validated.

The rather large tumor volumes (40±30cm^3) in our patients dataset might be considered as a limitation of this study, since partial volume effects are usually considered significant for volumes around or below 10 cm^3 (30). Firstly, 25% of the tumors in this dataset were within this volume range. In addition, it should be noted that the shape of the primary esophageal lesions is not spherical but mostly cylindrical with a small diameter (<2cm) in the transaxial direction. Therefore esophageal lesions can be significantly affected by partial volume effects despite the overall large metabolic volumes, as can be seen in figure 2 for a lesion with MATV above 25 cm^3. Finally, the patient population used in this study represents a typical clinical routine practice patient population and was not selected based on the overall primary MATVs.

CONCLUSIONS

The results of this study demonstrate that PVE correction does not add any value in parameters derived from metabolically active tumor volumes such as MATV and TLG measured on ^{18}F-FDG PET baseline acquisitions. PVC did not alter the already demonstrated clinical value of both parameters as predictive factors of the response to concomitant radio-chemotherapy or as prognostic factors of overall survival in locally advanced esophageal cancer. Similarly, although PVC led to increases in all SUV measurements (max, peak or mean) considered in clinical practice, the corrected values were still not significantly associated with either therapy response or prognosis. Finally, our study is in agreement with previous investigations using simpler tools, showing limited interest of PVC in this specific context. However, the potential impact of PVE correction in other applications such as diagnosis or lesion

detectability remains to be evaluated. In addition, the value of PVC in patient follow up using serial PET scans needs to be further demonstrated.

References

1. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin.* 2005;55:74-108.

2. Hayat MJ, Howlader N, Reichman ME, Edwards BK. Cancer statistics, trends, and multiple primary cancer analyses from the Surveillance, Epidemiology, and End Results (SEER) Program. *Oncologist.* 2007;12:20-37.

3. Gebski V, Burmeister B, Smithers BM, Foo K, Zalcberg J, Simes J. Survival benefits from neoadjuvant chemoradiotherapy or chemotherapy in oesophageal carcinoma: a meta-analysis. *Lancet Oncol.* 2007;8:226-234.

4. Kelsen DP, Winter KA, Gunderson LL, et al. Long-term results of RTOG trial 8911 (USA Intergroup 113): a random assignment trial comparison of chemotherapy followed by surgery compared with surgery alone for esophageal cancer. *J Clin Oncol.* 2007;25:3719-3725.

5. Stahl M, Wilke H, Stuschke M, et al. Clinical response to induction chemotherapy predicts local control and long-term survival in multimodal treatment of patients with locally advanced esophageal cancer. *J Cancer Res Clin Oncol.* 2005;131:67-72.

6. Aklilu M, Ilson DH. Targeted agents and esophageal cancer--the next step? *Semin Radiat Oncol.* 2007;17:62-69.

7. Hatt M, Visvikis D, Pradier O, Cheze-le Rest C. Baseline (18)F-FDG PET image-derived parameters for therapy response prediction in oesophageal cancer. *Eur J Nucl Med Mol Imaging.* 2011;38:1595-1606.

8. Larson SM, Erdi Y, Akhurst T, 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;2:159-171.

9. Hyun SH, Choi JY, Shim YM, et al. Prognostic value of metabolic tumor volume measured by 18F-fluorodeoxyglucose positron emission tomography in patients with esophageal carcinoma. *Ann Surg Oncol.* 2010;17:115-122.

10. Roedl JB, Prabhakar HB, Mueller PR, Colen RR, Blake MA. Prediction of metastatic disease and survival in patients with gastric and gastroesophageal junction tumors: the incremental value of PET-CT over PET and the clinical role of primary tumor volume measurements. *Acad Radiol.* 2009;16:218-226.

11. Mamede M, Abreu ELP, Oliva MR, Nose V, Mamon H, Gerbaudo VH. FDG-PET/CT tumor segmentation-derived indices of metabolic activity to assess response to neoadjuvant therapy and progression-free survival in esophageal cancer: correlation with histopathology results. *Am J Clin Oncol.* 2007;30:377-388.

12. Hatt M, Visvikis D, Albarghach NM, Tixier F, Pradier O, Cheze-le Rest C. Prognostic value of 18F-FDG PET image-based parameters in oesophageal cancer and impact of tumour delineation methodology. *Eur J Nucl Med Mol Imaging.* 2011;38:1191-1202.

13. Wahl RL, Jacene H, Kasamon Y, Lodge MA. From RECIST to PERCIST: Evolving Considerations for PET response criteria in solid tumors. *J Nucl Med.* 2009;50 Suppl 1:122S-150S.

14. Lee HY, Hyun SH, Lee KS, et al. Volume-based parameter of 18)F-FDG PET/CT in malignant pleural mesothelioma: prediction of therapeutic response and prognostic implications. *Ann Surg Oncol.* 2010;17:2787-2794.

15. Cazaentre T, Morschhauser F, Vermandel M, et al. Pre-therapy 18F-FDG PET quantitative parameters help in predicting the response to radioimmunotherapy in non-Hodgkin lymphoma. *Eur J Nucl Med Mol Imaging.* 2010;37:494-504.

16. Tixier F, Le Rest CC, Hatt M, et al. Intratumor heterogeneity characterized by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. *J Nucl Med.* 2011;52:369-378.

17. Kwee RM. Prediction of tumor response to neoadjuvant therapy in patients with esophageal cancer with use of 18F FDG PET: a systematic review. *Radiology.* 2010;254:707-717.

18. Hoetjes NJ, van Velden FH, Hoekstra OS, et al. Partial volume correction strategies for quantitative FDG PET in oncology. *Eur J Nucl Med Mol Imaging.* 2010;37:1679-1687.

19. van Heijl M, Omloo JM, van Berge Henegouwen MI, van Lanschot JJ, Sloof GW, Boellaard R. Influence of ROI definition, partial volume correction and SUV normalization on SUV-survival correlation in oesophageal cancer. *Nucl Med Commun.* 2010;31:652-658.

20. Hatt M, Cheze le Rest C, Descourt P, et al. Accurate automatic delineation of heterogeneous functional volumes in positron emission tomography for oncology applications. *Int J Radiat Oncol Biol Phys.* 2010;77:301-308.

21. Therasse P, Arbuck SG, Eisenhauer EA, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst.* 2000;92:205-216.

22. Boussion N, Cheze Le Rest C, Hatt M, Visvikis D. Incorporation of wavelet-based denoising in iterative deconvolution for partial volume correction in whole-body PET imaging. *Eur J Nucl Med Mol Imaging.* 2009;36:1064-1075.

23. Chang SG, Yu B, Vetterli M. Adaptive wavelet thresholding for image denoising and compression. *IEEE Trans Image Process.* 2000;9:1532-1546.

24. 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;28:881-893.

25. Hatt M, Cheze Le Rest C, Albarghach N, Pradier O, Visvikis D. PET functional volume delineation: a robustness and repeatability study. *Eur J Nucl Med Mol Imaging.* 2011;38:663-672.

26. Hatt M, Cheze-Le Rest C, Aboagye EO, et al. Reproducibility of 18F-FDG and 3'-deoxy-3'-18F-fluorothymidine PET tumor volume measurements. *J Nucl Med.* 2010;51:1368-1376.

27. Kruskal W, Wallis W. Use of ranks in one-criterion variance analysis. *J Am Stat Assoc.* 1952;47:583-621.

28. Kaplan E, Meyer P. Non parametric estimation from incomplete observation. *J Am Stat Assoc.* 1958;53:457-481.

29. Metz CE. Basic principles of ROC analysis. *Semin Nucl Med.* 1978;8:283-298.

30. Soret M, Bacharach SL, Buvat I. Partial-volume effect in PET tumor imaging. *J Nucl Med.* 2007;48:932-945.

31. Hatt M, Visvikis D. Defining radiotherapy target volumes using 18F-fluoro-deoxy-glucose positron emission tomography/computed tomography: still a Pandora's box?: in regard to Devic et al. (Int J Radiat Oncol Biol Phys 2010). *Int J Radiat Oncol Biol Phys.* 2010;78:1605.

32. Omloo JM, Sloof GW, Boellaard R, et al. Importance of fluorodeoxyglucose-positron emission tomography (FDG-PET) and endoscopic ultrasonography parameters in predicting survival following surgery for esophageal cancer. *Endoscopy.* 2008;40:464-471.

33. Chatterton BE, Ho Shon I, Baldey A, et al. Positron emission tomography changes management and prognostic stratification in patients with oesophageal cancer: results of a multicentre prospective study. *Eur J Nucl Med Mol Imaging.* 2009;36:354-361.

34. Hanin FX, Lonneux M, Cornet J, et al. Prognostic value of FDG uptake in early stage non-small cell lung cancer. *Eur J Cardiothorac Surg.* 2008;33:819-823.

Conflict of interest:

The authors have no conflict of interest.

Acknowledgments:

None.

Figures captions

Figure 1: Illustration of the iterative deconvolution partial volume effects correction on a whole-body ^{18}F-FDG PET image with (A) the original image and (B) the corrected image.

Figure 2: Qualitative differences between original and corrected PET images of an esophageal lesion of MATV above 25 cm^3 using profiles on axial, sagittal and coronal planes.

Figure 3: Examples of FLAB delineation results (blue contours) on the original (left) and the corrected (right) PET images with (A) a large slightly heterogeneous MATV, (B) a MATV with a necrotic core, (C) a small homogeneous MATV.

Figure 4: Examples of distributions of NR, PR and CR patients and associated Kruskal-Wallis tests results: (A) SUV_{max} and SUV_{max}^{PVC}, (B) MATV and $MATV^{PVC}$, (C) TLG and TLG^{PVC}.

Figure 5: Examples of Kaplan-Meier survival curves obtained using: (A) SUV_{peak}, (B) SUV_{peak}^{PVC}, (C) MATV.

Table captions

Table I: Patient demographic and clinical characteristics

Table II: Distributions of parameters with and without PVC.

Table III: Kruskal-Wallis test results (H statistic and associated *p* value) for each parameter, with the ability to differentiate ($p<0.05$) each pair of response group among patients (CR complete responders, PR partial responders and NR non responders).

Table IV: Univariate analysis results using Kaplan-Meier survival curves with the optimal threshold cutoff, associated hazard ratios (HR), 95% confidence intervals (CI) for HR, as well as the associated *p* value and median survival of each group.

Parameter	Number of patients (%)
Gender	
Male	45(90)
Female	5(10)
Age	
Range	45-84
Median	69
Site	
Upper esophagus	13(26)
Middle esophagus	20(40)
Lower esophagus	17(34)
Histology type	
Adenocarcinoma	14(28)
Squamous cell carcinoma	36(72)
Histology differentiation	
Well differentiated	14(28)
Moderately differentiated	12(24)
Poorly differentiated	5(10)
Unknown	19(38)
TNM Stage	
T1	7(14)
T2	8(16)
T3	24(48)
T4	11(22)
N0	20(40)
N1	30(60)
M0	34(68)
M1	16(32)
AJCC Stage	
I	4(8)
IIA	8(16)
IIB	6(12)
III	16(32)
IVA	16(32)

Table I

Definition	Notation	Original Mean ± SD	PVC Mean ± SD
Highest SUV	SUV_{max}	9.7 ± 3.9	14.9 ± 6.1
Mean of SUV_{max} and its 26 neighbors	SUV_{peak}	8.0 ± 3.3	10.1 ± 4.0
Mean SUV within MATV	SUV_{mean}	5.8 ± 2.4	7.4 ± 3.1
Metabolically active tumor volume (cm^3)	MATV	39.9 ± 36.1	36.2 ± 33.7
Total lesion glycolysis (g)	TLG	218.1 ± 208.3	235.8 ± 218.1

Table II

Parameter	H	*P*	Response differentiation? (p<0.05)		
			CR(n=11) / NR(n=14)	CR(n=11) / PR(n=25)	PR(n=25) / NR(n=14)
SUV_{max}	3.6	0.17	no	no	no
SUV_{max}^{PVC}	2.4	0.31	no	no	no
SUV_{peak}	5.1	0.08	no	no	no
SUV_{peak}^{PVC}	4.7	0.10	no	no	no
SUV_{mean}	5.5	0.07	no	no	no
SUV_{mean}^{PVC}	3.9	0.14	no	no	no
MATV	20.7	<0.0001	yes	yes	yes
$MATV^{PVC}$	20.7	<0.0001	yes	yes	yes
TLG	25.1	<0.0001	yes	yes	yes
TLG^{PVC}	25.2	<0.0001	yes	yes	yes

Table III

Parameter	Threshold	HR	HR 95%CI	*P*	Median survival (months)
SUV_{max}	8	1.5	0.7 to 3.1	0.28	20 vs. 13
SUV_{max}^{PVC}	11	1.6	0.7 to 3.2	0.26	20 vs. 13
SUV_{peak}	7	1.4	0.7 to 2.8	0.31	16 vs. 10
SUV_{peak}^{PVC}	9	1.8	0.9 to 3.6	0.11	20 vs. 11
SUV_{mean}	6.5	1.7	0.8 to 3.6	0.15	16 vs. 10
SUV_{mean}^{PVC}	7.5	1.7	0.8 to 3.5	0.12	20 vs. 10
MATV	85	3.9	1.0 to 15.2	0.0004	20 vs. 6
$MATV^{PVC}$	80	3.4	0.9 to 11.7	0.0024	16 vs. 10
TLG	260	2.9	1.2 to 6.8	0.0012	21 vs. 10
TLG^{PVC}	280	3.2	1.3 to 7.6	0.0004	21 vs. 10

Table IV

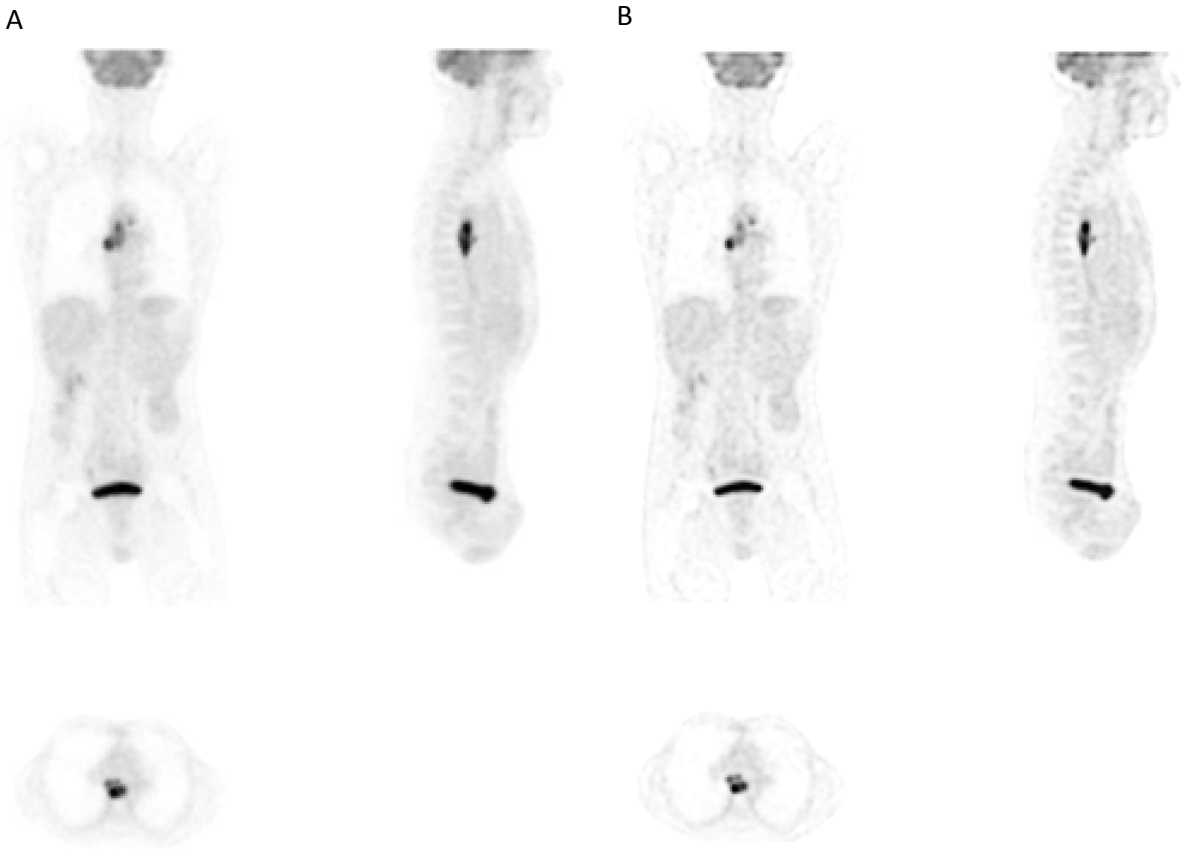


Figure 1

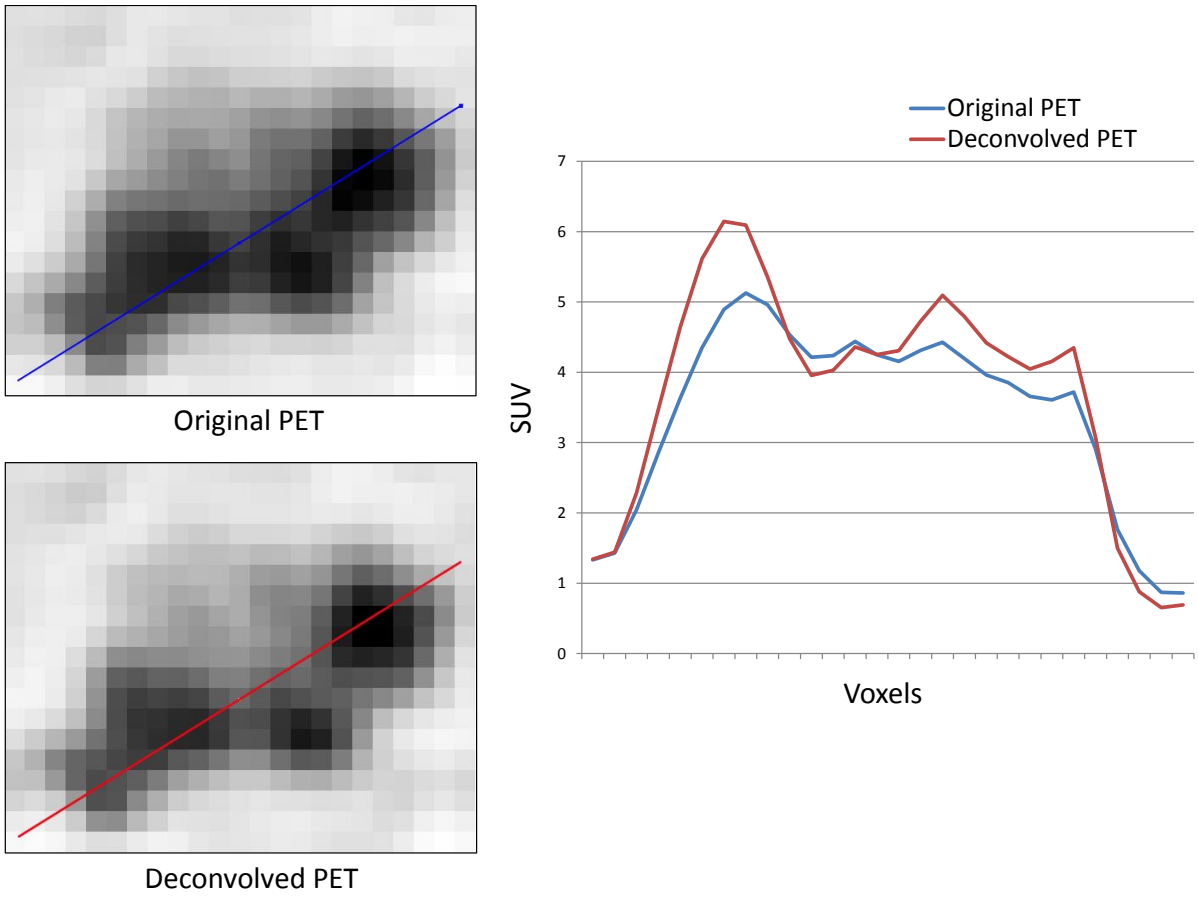

Original PET
Deconvolved PET
Original PET
Deconvolved PET
SUV
0
1
2
3
4
5
6
7
Voxels

Figure 2

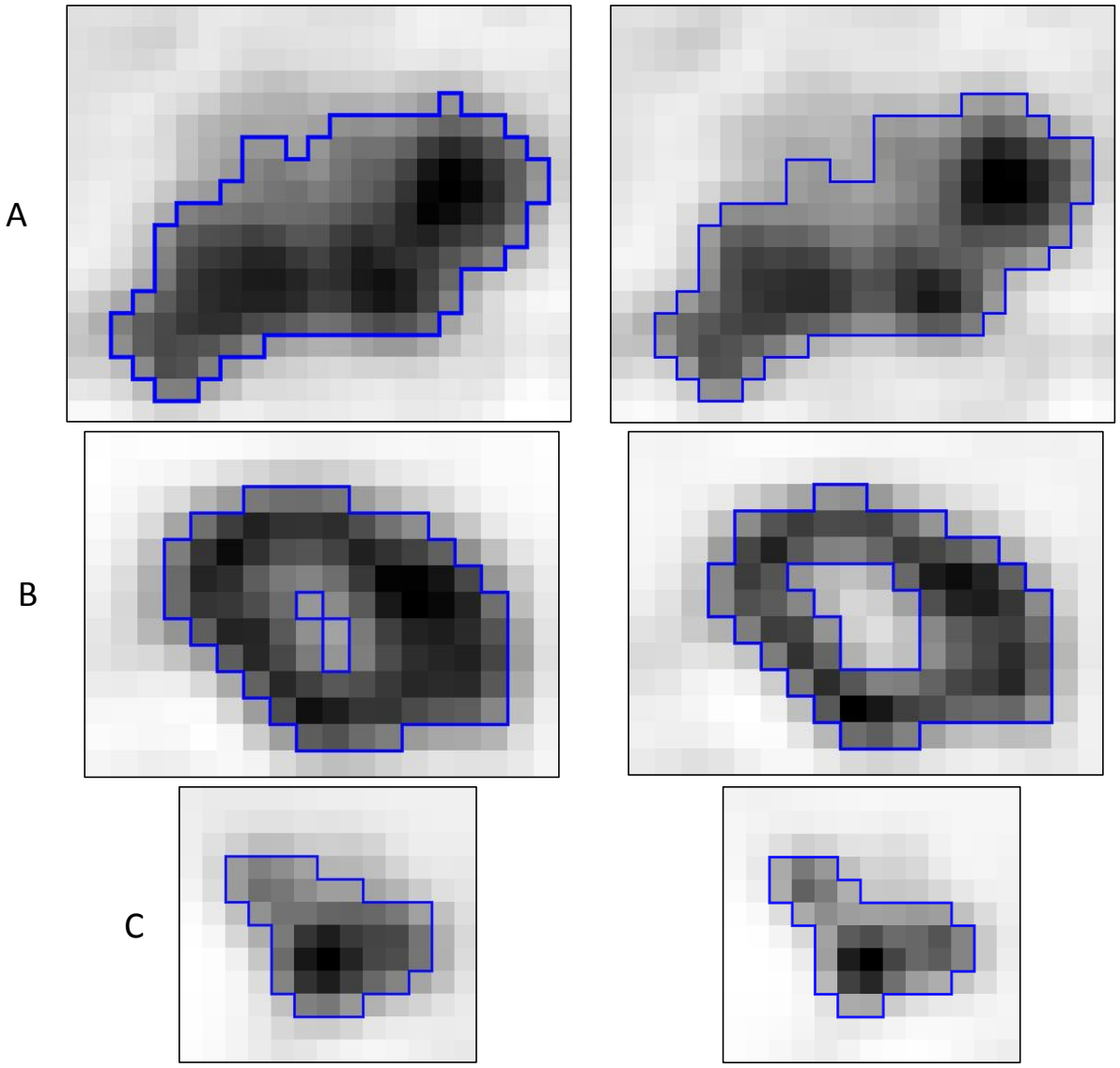
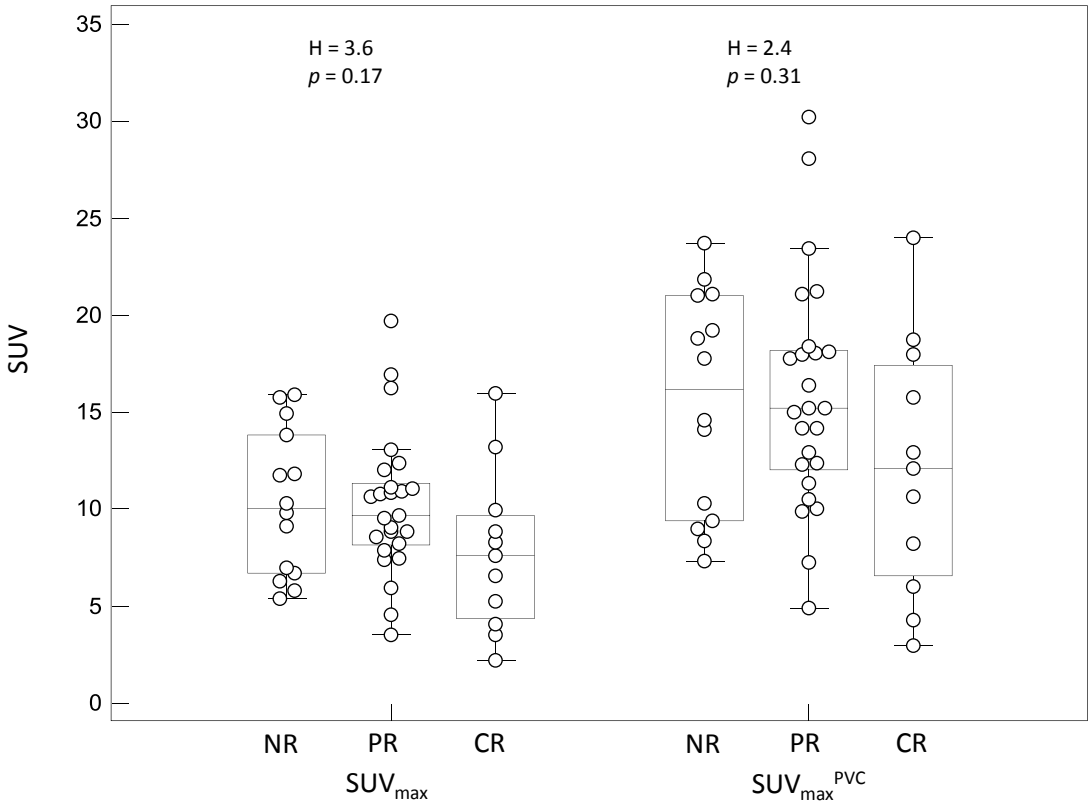
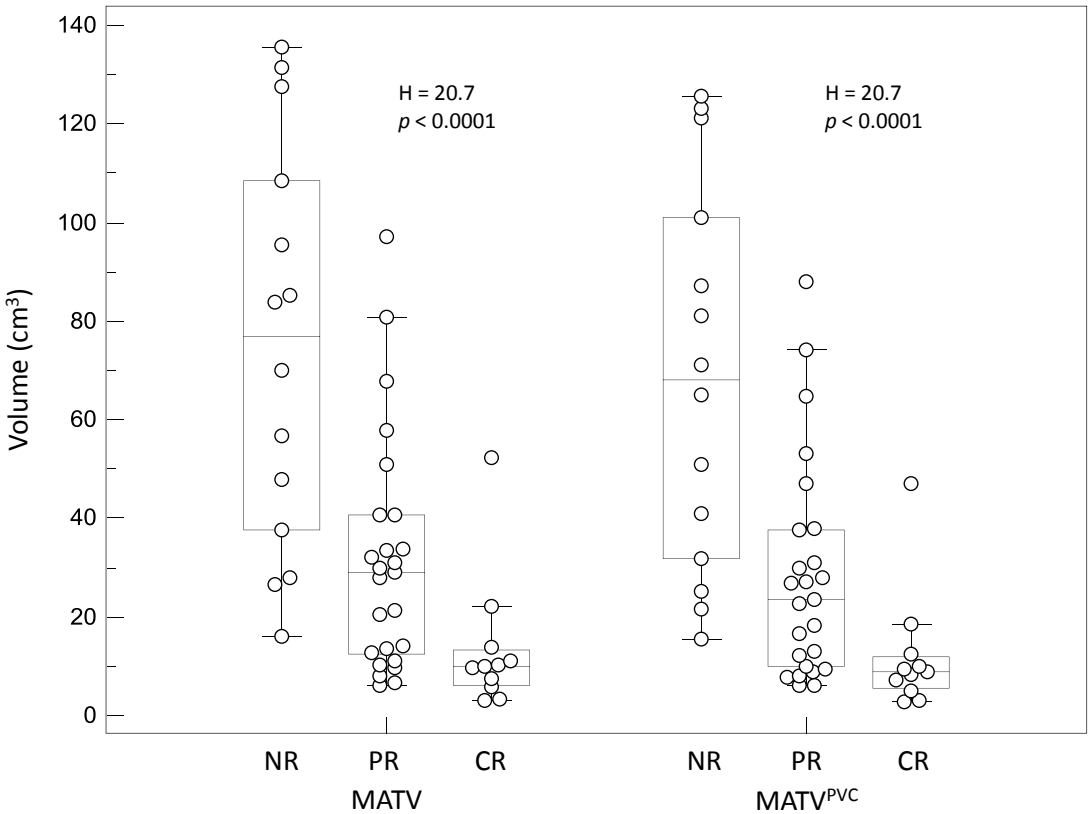


Figure 3

A


SUV
35
30
25
20
15
10
5
0
H = 3.6
p = 0.17
H = 2.4
p = 0.31
NR
PR
CR
NR
PR
CR
SUVmax
SUVmaxPVC

B


Volume (cm3)
140
120
100
80
60
40
20
0
H = 20.7
p < 0.0001
H = 20.7
p < 0.0001
NR
PR
CR
NR
PR
CR
MATV
MATVPVC

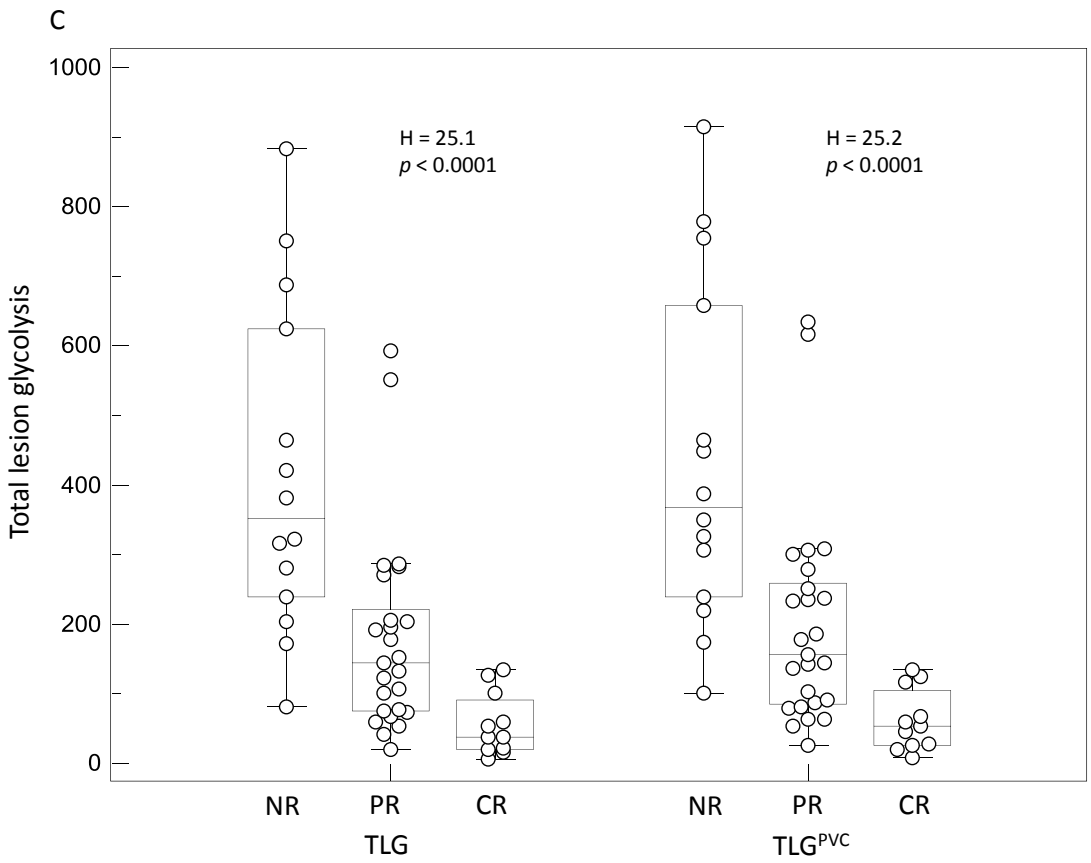
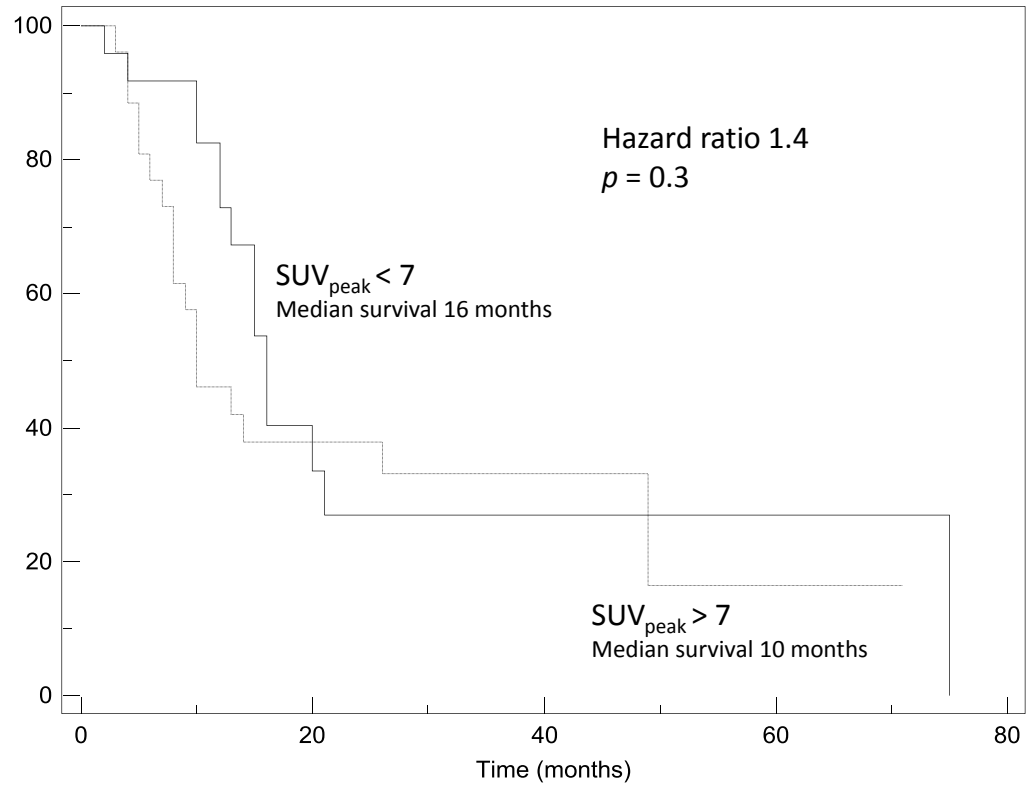
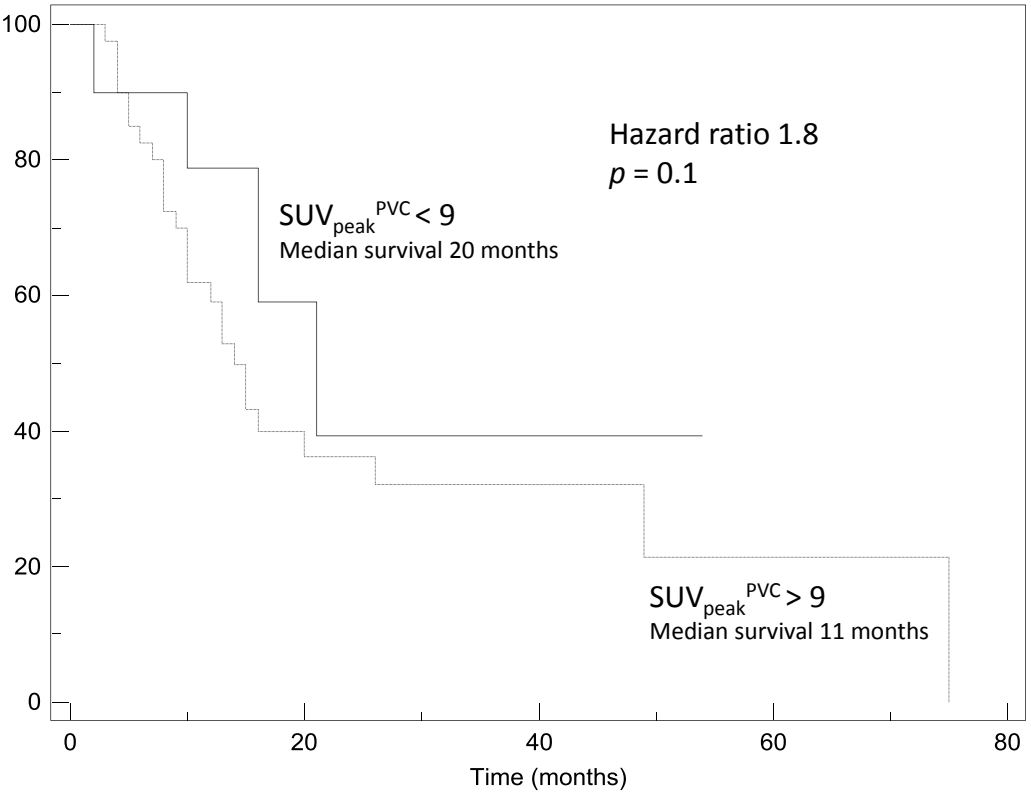

C
1000
800
600
400
200
0
Total lesion glycolysis
H = 25.1
p < 0.0001
H = 25.2
p < 0.0001
NR
PR
CR
TLG
NR
PR
CR
TLG^PVC

Figure 4

A


100
80
60
40
20
0
0
20
40
60
80
Time (months)
Hazard ratio 1.4
p = 0.3
SUVpeak < 7
Median survival 16 months
SUVpeak > 7
Median survival 10 months

B


100
80
60
40
20
0
0
20
40
60
80
Time (months)
Hazard ratio 1.8
p = 0.1
SUVpeak PVC < 9
Median survival 20 months
SUVpeak PVC > 9
Median survival 11 months

C

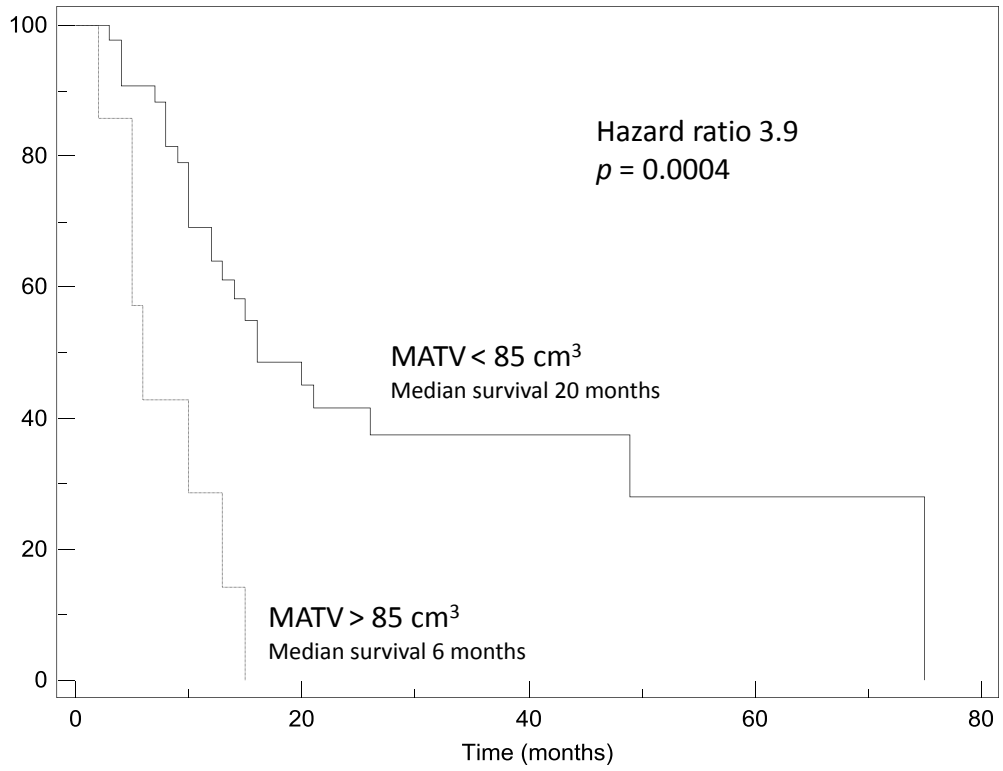

100
80
60
40
20
0
0
20
40
60
80
Time (months)
Hazard ratio 3.9
p = 0.0004
MATV < 85 cm3
Median survival 20 months
MATV > 85 cm3
Median survival 6 months

Figure 5