18F-FDG PET/CT heterogeneity quantification through textural features in the era of harmonisation programs: a focus on lung cancer

Charline Lasnon, Mohamed Majdoub, Brice Lavigne, Pascal Do, Jeannick Madelaine, Dimitris Visvikis, Mathieu Hatt, Nicolas Aide

HAL Id: hal-01386846
https://hal.science/hal-01386846

Submitted on 24 Oct 2016

^{18}F-FDG PET/CT heterogeneity quantification through textural features in the era of harmonization programs: focus on lung cancer.

Charline Lasnon[1-3], Mohamed Majdoub[4], Brice Lavigne[4], Pascal Do[5], Jeannick Madelaine[6],

Dimitris Visvikis[4], Mathieu Hatt[4*] and Nicolas Aide[1-3*]

[1]Nuclear Medicine Department, University Hospital, Caen, France
[2]INSERM UMR 1199, François Baclesse Cancer Centre, Caen, France
[3]Normandie University, Caen, France
[4]LaTIM, INSERM UMR 1101, Brest, France
[5]Thoracic oncology, François Baclesse Cancer Centre, Caen, France
[6]Pulmonology Department, Caen University Hospital, Caen, France

[*]Equally contributed to this work

Corresponding authors:

Prof Nicolas Aide

Nuclear Medicine Department
INSERM UMR 1199
Caen University Hospital
Avenue Côte de Nacre, 14000 Caen, France
Ph 0231063244 / Fax : 0231064927
E-mail : aide-n@chu-caen.fr

Dr Mathieu Hatt

INSERM UMR 1101, Laboratoire de Traitement de l'Information Medicale (LaTIM)
Groupe 'Imagerie multi-modalité quantitative pour le diagnostic et la thérapie'
CHRU Morvan, Bâtiment 1, 1er étage, bureau 1056,
2 avenue Foch, 29609 Brest, France
Ph: 02 98 01 81 11 / Fax: 02 98 01 81 24
E-mail: mathieu.hatt@inserm.fr

ABSTRACT

Purpose: Quantification of tumour heterogeneity in PET images has recently gained interest, but has been shown to be dependent on image reconstruction. This study aimed to evaluate the impact of the EANM/EARL accreditation program on selected ^{18}F-FDG heterogeneity metrics.

Methods: Seventy-one tumours in sixty biopsy-proven lung cancer patients acquisitions reconstructed with unfiltered PSF PET images (optimized for diagnostic purposes), PSF-reconstructed images with a 7mm Gaussian filter (PSF$_7$) chosen to meet EANM 1.0 harmonizing standards, and EARL-compliant OSEM images were prospectively included. Delineation was performed with FLAB algorithm on PSF images and reported on PSF$_7$ and OSEM ones, and with a 50% SUV$_{max}$ threshold applied independently to each image. Robust and reproducible heterogeneity metrics including 1st-order [area under the curve of the cumulative histogram (CH$_{AUC}$)], 2^{d}-order (entropy, correlation and dissimilarity) and 3rd-order [high intensity larger area emphasis (HILAE) and zone percentage (ZP)] textural features (TF) were statistically compared.

Results: Volumes obtained with 50% SUV$_{max}$ threshold were significantly smaller than FLAB-derived ones, and were significantly smaller in PSF images compared to OSEM and PSF$_7$ ones. PSF-reconstructed images showed significantly higher SUVmax and SUVmean values, as well as heterogeneity for CH$_{AUC}$, dissimilarity, correlation and HILAE, and a wider range of heterogeneity values than OSEM images for most of the metrics considered, especially when analysing larger tumours. Histological subtypes had no impact on TF distribution.

No significant difference was observed between any of the metrics considered (SUV or heterogeneity features) extracted from OSEM and PSF$_7$ reconstructions. Furthermore, the distributions of TF for OSEM and PSF$_7$ reconstructions according to tumour volumes were similar for all ranges of volumes.

Conclusion: PSF reconstruction with Gaussian filtering chosen to meet harmonizing standards resulted in similar SUV values and heterogeneity information, compared to OSEM images, which validates its use within the harmonization strategy context. However, unfiltered PSF-reconstructed images also showed higher heterogeneity according to some metrics, as well as a wider range of heterogeneity values than OSEM images, for most of the metrics considered, especially when analysing larger tumours. This suggests that, whenever

available, unfiltered PSF images should also be exploited to obtain the most discriminative quantitative heterogeneity features.

Introduction

Using quantitative parameters extracted from the Positron Emission Tomography (PET) component of PET/Computed Tomography (CT) images, such as Standardized Uptake Values (SUV), as biomarkers in multicentre trials or in sites equipped with multiple PET/CT scanners requires that these parameters be comparable among patients, regardless of the PET/CT system used. This can be achieved by harmonizing patient preparation, as well as data acquisition, reconstruction and processing, including the image analysis and parameters extraction steps [1-3]. The American college of Radiology (ACR) program [4], the EANM/EARL accreditation program [5] and the SNM clinical trials network (SNM-CTN) [6] have set up harmonization programs based on the use of phantoms acquisitions. These are used as standardized objects in order to harmonize data acquisition, processing and analysis so that the physical, technical and biological sources of error [1, 7] in SUV measurements can be limited.

A specific issue is related to reconstruction-dependent variations encountered with recently introduced advanced image reconstruction algorithms such as those incorporating the point spread function (PSF) in the system matrix [8], or Bayesian penalised likelihood (BPL) reconstruction [9]. These new image reconstruction schemes have been shown to produce SUV metrics significantly higher than conventional ordered subset expectation maximization (OSEM) algorithms [10]. Consequently, an additional filtering step can be used in order to meet harmonizing standards [11-13]. With regards to the EANM/EARL program [5], a set of PET images with NEMA NU-2 anthropomorphic phantom-based filtering is mandatory to harmonize SUVs to the EANM standards. Given that centres running PET systems with advanced reconstruction algorithms are often willing to use them with parameters chosen in order to achieve optimal lesion detection, EARL-accredited centres tend to use two PET

datasets when participating in multicentre trials: one for optimal lesion detection and image interpretation, and the filtered one for harmonized quantification [12].

However, it is important to emphasize that all these previous efforts have been focused on usual SUV metrics, as they are commonly used in oncology for therapy assessment and risk stratification. However, there is growing interest in using alternative measurements, for instance metabolically active tumour volume (MATV) and heterogeneity metrics, in order to provide a more comprehensive quantitative assessment of lesions from PET images [14, 15]. One of the most promising approaches for heterogeneity quantification is textural features analysis, introduced for image processing applications in the 70's, used in MRI and CT since the early 90's and more recently in PET [16]. As PSF reconstruction improves resolution and therefore provides higher definition of structures within a lesion, it is reasonable to expect improved evaluation of tumour heterogeneity as compared to OSEM algorithms. This raises the question of which reconstruction should be used for assessing tumour heterogeneity within a program using a smoothed dataset to reach harmonizing standards. Two studies have already reported on the impact of the type of reconstruction algorithm or variation of reconstruction parameters on the textural features values [17, 18]. However, they have mostly been focused on reporting the quantitative impact only, and have neither explored the issue within the context of harmonization programs nor the relationship between heterogeneity and volume, an important aspect recently demonstrated [19].

This study focused on lung cancer, a tumour type for which standard SUV metrics have been proved to be clinically useful [20-23] and for which quantification of tumour heterogeneity in PET images has recently gained interest [24, 25] and aimed at evaluating the potential impact of the EARL accreditation program [5] on selected ^{18}F-FDG heterogeneity metrics. The primary aim was to compare several heterogeneity features previously identified as reliable (robust and repeatable) in lung cancer patients, in PSF-reconstructed images, PSF-

reconstructed images with a filter chosen to meet harmonizing standards and in EARL compliant OSEM images, later referred as to OSEM images. This comparison was performed not only in terms of absolute values but also in terms of their distributions with respect to tumour volume, which was not considered in previous studies. A secondary aim was to study whether potential differences in heterogeneity features amongst these three reconstructions would be similar in adenocarcinomas (ADC), squamous cell carcinomas (SqCC), and large cell lung cancer (LCC), the main histological types encountered in non-small cell lung cancers

Materials and methods

Patients' selection

Over a 3-month period, 60 consecutive biopsy proven lung cancer patients (4 small-cell lung cancer and 56 NSCLC) were prospectively included. Informed consent was waived for this type of study by the local ethics committees (Ref A12-D24-VOL13, *Comité de protection des personnes Nord-Ouest III*) since the scans were performed for clinical indications and the study procedures were performed independently of usual clinical reporting.

PET calibration and cross calibration

The calibration of the PET system was performed daily with a ^{68}Ge cylinder with a known radioactive concentration.

The cross-calibration procedure was performed twice during the present study, as per the EANM guidelines [11]. Details regarding this cross-calibration can be found elsewhere [12]. The cross-calibration factors were found to be 0.99 and 1.00.

PET/CT examinations

After a 15-min rest in a warm room, patients who had been fasting for 6 h were injected with ^{18}F-FDG. The injected activity and the exact delay between injection and the start of the acquisition were recorded for each patient.

All PET imaging studies were performed on a Biograph TrueV (Siemens Medical Solutions) with a 6-slice spiral CT component. For additional technical details regarding this system we refer to a previous publication [26]. CT acquisition was performed first, with the following parameters: 60 mAs, 130 kVp, pitch 1 and 6×2 mm collimation. Subsequently, the PET

emission acquisition was performed in 3-D mode. Patients were scanned from the skull base to the mid-thighs.

PET reconstruction

The standard reconstruction in the department where patients were recruited is a PSF reconstruction algorithm (HD; TrueX, Siemens Medical Solutions; 3 iterations and 21 subsets) without filtering. For the purpose of the present study, raw data were also reconstructed with the OSEM reconstruction algorithm (4 iterations and 8 subsets) and a PSF reconstruction algorithm (HD; TrueX, Siemens Medical Solutions; 3 iterations and 21 subsets) with a 7 mm Gaussian filter (PSF_7). As shown in a previous study, this latter reconstruction leads to protocol-specific images with NEMA NU-2 phantom-based filtering that meet EANM quantitative harmonizing standards, therefore reducing reconstruction-dependent variation in SUVs [12]. The OSEM reconstruction parameters met the EANM requirements regarding activity recovery.

For all reconstructions, matrix size was 168×168 voxels, resulting in isotropic voxels of $4.07 \times 4.07 \times 4.07$ mm^3. Scatter and attenuation (using the associated CT) corrections were applied.

PET tumour delineation

All lesions were first automatically delineated in the PSF PET image using the Fuzzy Locally Adaptive Bayesian (FLAB) algorithm, and resulting segmentations were reported on the two other images (OSEM and PSF_7). This process avoided any variability in the tumour volumes definition and number of voxels involved in the calculations when comparing features across the three images. FLAB has been developed specifically for PET image segmentation [27]

and has been thoroughly validated for reproducibility, robustness and repeatability [28, 29], as well as for accuracy on simulated and clinical images [30]. Tumours were first located and isolated in a volume of interest (VOI) well enclosing the tumour and its surrounding background, without including nearby pathological uptake. This was performed using in-house software in which points are placed by the user around the tumour (see supplemental Fig. 1). The FLAB algorithm was then applied to this VOI in fully automated mode, in contrast to a semi-supervised approach considered in a previous work on FLT images during radiotherapy, in which the contrast and signal-to-noise ratio were lower [31].

However, in order to be more representative of a current multicentric clinical setting, tumour volumes were also defined using a fixed threshold at 50% of SUV_{max} (T50) applied independently to each of the three images. Furthermore, for the most discordant volumes between PSF7/OSEM and PSF (outliers located above the 90th percentile) when using T50, tumours were also segmented independently with FLAB on each image.

Features were then extracted from each of the three volumes and compared.

Tumour characterization and quantification

From the FLAB-delineated volumes, all "standard" metrics were extracted: SUV_{max}, SUV_{mean} and metabolically active tumour volume. To characterize uptake heterogeneity, several metrics were considered: on the one hand, a first-order metric based on the intensity histogram (IH) [16] denoted area under the curve of the cumulative histogram (CH_{AUC}) [32]. On the other hand, second- and third-order textural features (TF). The metric CH_{AUC} is based on intensity histogram only and does not incorporate spatial information. TF have been defined to quantify patterns of spatial arrangements and/or intensity variations. There exist dozens of TFs based on different computational frameworks. In the present work, we used only a few selected TFs. This selection was based on several previous studies showing that

most of the features (including first-order metrics such as skewness), especially third-order metrics focusing on small areas and/or low intensities, are unreliable due to poor robustness *vs.* reconstruction [17, 18] or partial volume effects and segmentation [33], and low reproducibility on test-retest images [34]. The remaining features are either calculated from co-occurrence (2nd order, entropy, correlation and dissimilarity) or size-zone matrices [3rd order, high intensity larger area emphasis (HILAE) and zone percentage (ZP)]. Before building these matrices, images are first discretized into a chosen number of bins (B) with a quantization step. It has been shown that the choice of the quantization value (usually between 8 and 256) has an important impact on the resulting TF value, but also reproducibility [34] or complementary value with the tumour volume in which it is calculated [19]. Based on these previous results, a value of B=64 was used in the present work, and the quantization was performed using equation 1 in which *I(x)* is the original SUV of the voxel of interest and SUV*min* and SUV*max* are the minimal and maximal SUV value within the tumour volume.

$$I_B(x) = B \times \frac{I(x) - SUV_{min}}{SUV_{max} - SUV_{min}}$$

For co-occurrence matrices, it has been shown that less redundant features are obtained when calculated using a single co-occurrence matrix taking into account all 13 spatial directions simultaneously, rather than computing a matrix for each direction followed by averaging [19, 35]. A single co-occurrence matrix was thus adopted in the present work.

Noise analysis in PET images

In order to evaluate noise characteristics of each of the three reconstructions, signal-to-noise ratio (SNR, defined as $20 \times \log_{10}\left(\frac{\mu}{\sigma}\right) DB$ [36] where μ and σ are the mean and standard

deviation of intensities) was measured in circular regions of interest placed in homogeneous regions of the liver and automatically reported in each reconstruction.

Statistical analysis

Quantitative data are presented as mean (standard deviation), as well as the median when not normally distributed. Bland-Altman analyses were used to compare the SUV metrics obtained in the three images. The features obtained on each of the three sets of PET images were first compared globally using Friedman tests. Graphical plots of each feature depending on tumour volume were also used to estimate the impact of PSF reconstruction compared to OSEM and PSF_7 images, and the features were then compared by categories of volumes using Friedman tests. MATV, SUV_{max} and TFs extracted from the three sets of data were compared according to the histological type of the tumour (ADC, ScCC and LCC) using Kruskal-Wallis tests. For all tests, a two-tailed P value of less than 0.05 was considered statistically significant. Graphs and analyses were carried out using Prism (GraphPad Software, La Jolla, CA).

Results:

Population characteristics and compliance to guidelines for tumour imaging

Population characteristics are displayed in table 1. Overall, 58 (96.7%) patient examinations fulfilled the EANM 2.0 guidelines for PET tumour imaging. The mean (SD) injected dose of ^{18}F-FDG was 4.02 (0.16) MBq/kg. The mean (SD) delay between the injection and the start of the PET acquisition was 60.43 (3.38) min. The mean (SD) blood glucose level was 1.04 (0.23) mmol/L.

Validation of the use of an additional harmonized PET dataset to overcome reconstruction-dependency of SUVs

Overall, 71 pulmonary lesions were delineated. The mean (SD, median) FLAB-derived MATV was 31.7 (46.4, 9.7) cm^3. The mean (SD) SUV$_{max}$ for OSEM, PSF and PSF$_7$ reconstructions were 10.50 (5.85), 15.42 (9.56) and 10.56 (5.88) respectively. The mean (SD) SUV$_{mean}$ for OSEM, PSF and PSF$_7$ reconstructions were 6.14 (2.99), 7.37 (4.03) and 6.25 (2.98) respectively.

As shown in the supplemental Fig. 2, Bland-Altman analysis demonstrated that the mean ratio of PSF and OSEM reconstructions for SUV$_{max}$ and SUV$_{mean}$ were 1.46 (95%CI $= 0.86 - 2.08$) and 1.19 (95%CI $= 0.71 - 1.67$) respectively. When using the PSF$_7$ harmonized reconstruction, the mean ratio between PSF$_7$ and OSEM reconstructions were 1.01 (95%CI $= 0.93 - 1.09$) and 1.02 (95%CI $= 0.95 - 1.09$) for SUV$_{max}$ and SUV$_{mean}$ respectively.

Compared to OSEM, SNR in the liver was lower in PSF images (-25.8 ± 3.9 %) whereas it was very similar in PSF$_7$ images (1.2 ± 2.9 %).

Impact of newest reconstruction algorithms on textural features

In a first step we used FLAB to delineate lesions in PSF images and we reported this segmentation on OSEM images. For the first-order metric CH_{AUC} based on the intensity histogram, most tumours were quantified as significantly more heterogeneous in PSF images compared to OSEM images, as PSF values were significantly lower than OSEM ones (a lower area under the curve indicating higher heterogeneity). Regarding second-order metrics calculated on the co-occurrence matrix (entropy, correlation and dissimilarity), on the one hand, PSF values were significantly lower for correlation and significantly higher for dissimilarity (in both cases indicating higher heterogeneity), compared to OSEM reconstruction. On the other hand, no significant difference between PSF and OSEM images was observed for entropy. Regarding third-order metrics calculated on size-zone matrices (HILAE and ZP), there was a significant difference between PSF and OSEM images only for HILAE values, which were lower, indicating higher heterogeneity. Fig. 1 displays TFs for the three reconstructions used.

Heterogeneity features were also analysed depending on the range of tumour volumes. As shown in Fig. 2, the dispersion of the values (represented by the interquartile range) was larger and calculated values were significantly smaller for PSF reconstruction as compared to OSEM reconstruction, for tumour volumes larger than 1 cm^3 in the case of CH_{AUC}. For HILAE, calculated values were also significantly smaller for tumour volumes larger than 1 cm^3 but the dispersion of values was narrower for PSF compared to OSEM. Dissimilarity values were significantly higher in PSF for volumes > 50 cm^3 and the dispersion of these values was larger in PSF images for tumour volumes larger than 1 cm^3, compared to OSEM reconstruction. Distributions for all metrics can be seen in details in the supplemental Fig. 3 and Fig. 4.

When defining volumes using T50 applied independently to OSEM and PSF images, mean (median, SD) MATV were significantly smaller on PSF images [9.0 (12.4, 3.3) cm^3] as

compared to OSEM ones [18.8 (24.2, 7.3) cm^3] (p<0.0001). There were significant differences between PSF and OSEM images for first-order, second-order and third-order metrics with the same trends above detailed (Fig. 3). Fig. 4 displays representative examples of tumour delineation using the FLAB algorithm and a 50% of SUV$_{max}$ threshold, as well as the related SUV and TF metrics.

When comparing volumes obtained with T50 on PSF to OSEM and PSF$_7$ ones, 8 outliers (above the 90th percentiles) were observed and re-processed by independently contouring with FLAB. There was no significant difference between PSF and OSEM mean (median, SD) MATV [12.8 (12.2, 4.8) cm^3 and 13.4 (13.0, 4.8), respectively (+5.3±4.6%)]. There were significant differences between PSF and OSEM images for first-order, second-order and third-order metrics as previously described, except for dissimilarity and ZP for which there were only trends (Supplemental Fig. 5).

Effect of the harmonization strategy on textural features

In the previous section, using the FLAB algorithm to delineate lesions, we found that data extracted from OSEM reconstruction were different from those extracted from PSF reconstruction for several TFs but also for CH$_{AUC}$. When comparing these metrics extracted from OSEM and PSF$_7$ reconstructions, none exhibited significant differences (Fig. 1). Furthermore, the distributions of their values according to MATV were much more similar for all ranges of volumes (Fig. 2) with no significant difference whatever range of tumour volumes was considered (except for HILAE in volumes larger than 50 cm^3), highlighting that the quantifiable heterogeneity content of the PSF$_7$ images was very close to the one contained in OSEM images.

We also defined tumour volumes using T50 applied independently to OSEM and PSF$_7$ images. When using this methodology, no significant difference was observed between

MATV, CH_{AUC} and all TFs extracted from OSEM and PSF_7 reconstructions (Fig. 3). Mean (SD, median) MATV were 18.8 (24.2, 7.3) cm^3 and 19.5 (25.5, 7.7) cm^3 for OSEM and PSF_7 reconstruction respectively (ns).

Analyzing the 8 outliers described above for which FLAB was used independently on the three sets of images, there was no difference between OSEM and PSF_7 mean (median, SD) MATV [13.4 (13.0, 4.8) cm^3 and 13.6 (13.2, 5.11) respectively, $+1.0\pm3.4\%$] and between textural features extracted from OSEM and PSF_7 images (Supplemental Fig. 5).

SUV metrics and heterogeneity features amongst the histological subtypes

Standard metrics exhibited significant differences amongst the three NSCLC histological subtypes. In particular, there was a trend towards smaller volumes in ADC. SUV_{max} values were also different in the three subtypes, with however large overlaps between the three distributions. SUV_{max} values obtained in PSF reconstructed images were higher for all three subtypes, resulting however in a similar and unchanged differentiation between them: ADC had significantly lower SUV_{max} than SqCC and LCC in all three reconstructions (Fig. 5).

Although the heterogeneity metrics were differently distributed with the three different reconstruction schemes, none of them were significantly different among the three histological subtypes, whatever reconstructed image set was considered (Fig. 6).

Discussion:

Heterogeneity metrics, especially textural features, have gained interest in the past few years to quantify intratumour heterogeneity in PET images. There have been several studies highlighting the dependency of these metrics to various factors, including the image analysis workflow (such as tumour delineation or partial volume effects correction) [33], the image reconstruction schemes or parameters [17, 18], and basic stochastic effects occurring in the PET acquisition process [37].

Our results confirm some of these previous results regarding the impact of the reconstruction choices on these metrics values [18]. Compared to OSEM images, unfiltered PSF-reconstructed images showed lower signal-to-noise ratio in the liver, higher heterogeneity and higher range of heterogeneity values in the tumour, for most of the metrics when using FLAB (independently on the three sets of images or not) and for all of the metrics considered in the present work when using T50, to be more representative of a current multicentric clinical setting. This difference was logically especially observed when analysing larger tumours. Our study indeed sheds light on the impact of reconstruction algorithms on the distributions of heterogeneity features with respect to tumour volume, which had not been considered in these previous studies. Regarding the differences observed in the case of T50, it should be emphasized that part of these can be directly attributed to the fact that this segmentation method applied to PSF images led to significantly smaller volumes than on OSEM and PSF$_7$ images, with sometimes drastically reduced volumes not covering the tumour uptake spatial extent (see Fig. 4c and 4d and supplemental Fig. 6). This method which has been evaluated previously mostly on standard non-PSF images is thus clearly not appropriate to extract tumour volume and associated metrics from PSF-reconstructed images because of their higher contrast. This also strengthens our approach of using a single volume for more accurate and

rigorous comparison of the three images content in terms of textural patterns, to avoid any bias in the calculation of TF due to different volumes.

Thus, the impact of reconstruction for comparable tumour volumes was found to be significant for some metrics (CH_{AUC}, correlation, dissimilarity, HILAE) and only an observable trend for others (entropy, ZP), and the differences increased with larger tumour volumes (for instance in the case of ZP, differences were significant only for tumours larger than 50 cm^3). This suggests that PSF-based reconstruction may provide more quantifiable heterogeneity-related information in larger tumours than OSEM images, as the interval between smallest and highest values increases, thereby providing more potential for differentiating different levels of heterogeneity in these tumours. Our results also highlight the fact that some TFs seem more sensitive than others to the changes in PSF-reconstructed images compared to OSEM images when analysing similar volumes determined with FLAB: CH_{AUC}, correlation and HILAE showed higher sensitivity with larger differences in both overall and volume-related distributions than entropy, dissimilarity and ZP.

The present study was conducted within the overall harmonization strategy context and focused on the EARL accreditation program, which is why unfiltered PSF images (optimized for diagnostic purposes) and OSEM images were compared to PSF images filtered with a 7 mm Gaussian filter chosen to meet the EANM 1.0 guidelines (PSF_7). As previously published [12], the use of PSF_7 resulted in signal-to-noise ratio in the liver and SUV_{max} values in the tumour very close to OSEM, and the same pattern was observed for heterogeneity metrics. All metrics considered in the present study were very close with no significant differences when extracted from OSEM and PSF_7 images, whatever the delineation technique used. This suggests that OSEM and PSF_7 EARL compliant images present a similar quantifiable heterogeneity content, and validates the use of TFs extracted from PSF-filtered images for multicentre studies. However, as stated above, our results also suggest that using unfiltered

PSF-based reconstructions could provide potentially more discriminative image features allowing for higher differentiation amongst patients, for studies aiming to quantify tumour heterogeneity using TFs and exploit these metrics for a clinical endpoint, such as patient stratification according to survival or response to therapy. Of note, these studies have been mostly performed in single sites, but future validation studies will likely require pooling data from several centres in order to obtain larger cohorts with enough statistical power. This raises the issue of using filtered-harmonized PSF images so that they can be pooled with OSEM data from other centres (potentially losing some discriminative power from TFs), or pooling data only from centres using PSF reconstruction with no post-filtering. This issue is problematic, as the EARL accreditation program is not meant to exclude images from centres running PET systems not equipped with PSF reconstruction or other advanced algorithms. Also, the sensitivity of TFs to reconstruction parameters needs to be interpreted in the context of an important reconstruction disparity within PET centres, even in centres running the same PET system, as recently reported by the Clinical Trials Network of the SNMMI [38]. Taken together, these findings suggest that PSF reconstruction with a Gaussian filtering chosen to meet harmonizing standards could be used within the harmonization strategy context for studies aiming to quantify intra-tumour heterogeneity to stratify, rank or classify patients with respect to a given clinical endpoint. In addition, we recommend that whenever available, unfiltered PSF images should also be analysed, especially for large, single-centre series since quantitative metrics obtained from these could potentially offer higher discriminative power. This of course requires additional studies in larger cohorts to be demonstrated.

One limitation of our study is the inclusion of a single system where the underlying reconstruction was identical apart from the use of the PSF modeling. Differences in image reconstruction methods between several vendors may give rise to additional variability that needs to be evaluated before texture analysis could be reliably used in the context of

multicentric studies, despite the demonstrated repeatability and robustness of several features versus changes in image properties [17, 18, 39].

Finally, we sought to identify differences in heterogeneity features within NSCLC histological subtypes. Although we showed that the ADC presented much lower volumes as well as lower SUV_{max} values than the SqCC and LCC subtypes, none of the heterogeneity metrics showed any discriminative power in differentiating these subtypes, in either reconstruction method used, which is in line with recent results obtained in breast cancer [40]. On the other hand, it contradicts another recent study that suggested textural features could differentiate between ADC and SqCC in a cohort of 30 NSCLC Asian patients [41]. These results were obtained on 2D-slice –not 3D volume analysis only, and required the combination of numerous parameters through machine learning (automated clustering) in order to differentiate the two subtypes. The derived model was not validated in an external cohort. This possibly led to overfitting and the results might not be generalizable to other series of patients, especially European ones. Our results suggest that heterogeneity features could be used in a multicentre setting regardless of the histology in series of European patients. Indeed, results in term of heterogeneity in Asian patients may not be applicable to European patients, as not only the ratio between ADC and SqCC is inverted in these populations, but also the rate of EGFR mutation is higher in Asian patients (20-40%) compared to European patients (around 10%) [42]. One could therefore postulate different TFs in ADC depending on the mutation status. Studies with a larger ADC population, focusing on the differences in heterogeneity features between mutated and non-mutated ADC are therefore required to complement recent data on standard SUV metrics [43, 44].

Conclusion:

The use of PSF reconstruction with Gaussian filtering chosen to meet harmonizing standards produced comparable SUV values, as well as similar levels of heterogeneity information, compared to OSEM images, which validates its use within the harmonization strategy context for studies aiming to quantify intra-tumour heterogeneity to stratify, rank or classify patients. However, unfiltered PSF-reconstructed images showed significantly higher heterogeneity for CH_{AUC}, correlation and HILAE, as well as a wider range of heterogeneity values than OSEM ones, for most of the metrics considered, especially when analysing larger tumours. This suggests that, whenever available, unfiltered PSF images should also be analysed because resulting quantitative heterogeneity features could be more discriminative in stratifying or ranking patients, which remains to be demonstrated. Finally, the main NSCLC histological subtypes in this cohort did not show any differences in terms of intra-tumour heterogeneity, despite some notable differences in metabolically active tumour volume and levels of uptake (SUV_{max}). This may facilitate the potential multicentre use of heterogeneity features regardless of the histology in series of European patients.

Compliance with ethical standards

Conflicts of interest none to declare

Ethical approval All procedures performed in studies involving human participants 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.

Informed consent Informed consent was waived for this type of study by the local ethics committee (Ref A12-D24-VOL13, Comité de protection des personnes Nord-Ouest III), since the PET scans were performed for clinical indications and the trial procedures were performed independent of usual clinical reporting.

Table 1: Patient demographics

Characteristic	
Sex ratio (M/F)	3.13
Age (years)	
Range	37 – 90
Mean (SD)	64 (10.7)
Body habitus, n (%)	
BMI≤25	38 (63.3)
BMI>25	22 (36.7)
Histological diagnosis, n (%)	
Small cell lung cancer	4 (6.7)
Non-small cell lung cancer	56 (93.3)
Squamous cell carcinoma	28 (50.0)
Adenocarcinoma	21 (37.5)
Large cell lung cancer	7 (12.5)
Per patient lesions, mean (SD)	1.18 (0.46)

Figure legends

Fig. 1: Impact of the EARL harmonization strategy on textural features using the FLAB algorithm to delineate lesions

Textural features for the three reconstructions used. CH_{AUC}: area under the curve of the cumulative histogram; high intensity larger area emphasis (HILAE); ZP: zone percentage. Data is shown as Tukey boxplots (lines displaying median, 25th and 75th percentiles; cross represents the mean value).*, **, and *** indicate two-tailed $P < .05$, $P < .01$, and $P < .001$, respectively. ns : non significant

Fig. 2: Impact of tumour volume on textural features

Textural features used for the three reconstructions, depending on tumour volume. CH_{AUC}: area under the curve of the cumulative histogram; high intensity larger area emphasis (HILAE); ZP: zone percentage. Data is shown as Tukey boxplots (lines displaying median, 25th and 75th percentiles; cross represents the mean value).*, **, and *** indicate two-tailed $P < .05$, $P < .01$, and $P < .001$, respectively. ns: non significant

Fig. 3: Impact of the EARL harmonization strategy on textural features using a 50% of SUV_{max} threshold to delineate lesions

Textural features for the three reconstructions used. CH_{AUC}: area under the curve of the cumulative histogram; high intensity larger area emphasis (HILAE); ZP: zone percentage. Data is shown as Tukey boxplots (lines displaying median, 25th and 75th percentiles; cross represents the mean value).*, **, and *** indicate two-tailed $P < .05$, $P < .01$, and $P < .001$, respectively.

Fig. 4: Representative examples of lung tumours.

OSEM, PSF and PSF_7 images and textural features are displayed for a 67-year old male patient with a squamous cell carcinoma (panels a and c) and for a 44-year old female patient with an adenocarcinoma (panels b and d). Images have been scaled on the same maximum value. Note the improvement in tumour apparent activity and contrast in PSF images compared to OSEM images, and the similarity between OSEM and PSF_7 images. This can also be observed quantitatively in the extracted SUV and TF metrics. Green contours (panels a and b) denote the tumour delineation using the automatic Fuzzy Locally Adaptive Bayesian (FLAB) algorithm on the PSF image and reported on the two other datasets. Red contours (panels c and d) denote the delineation using the 50% of SUV_{max} threshold applied independently to each image. CHAUC: area under the curve of the cumulative histogram; high intensity larger area emphasis (HILAE); ZP: zone percentage.

Fig. 5: Standard quantification metrics

FLAB-derived metabolically active tumour volume (MATV) and standardized uptake values (SUVs) according to the histological subtypes (adenocarcinoma:ADC; squamous cell carcinoma:SqCC; large cell carcinoma:LCC) in non-small cell lung cancer patients for the three reconstructions used. Data is shown as Tukey boxplots (lines displaying median, 25th and 75th percentiles; cross represents the mean value). * and ** indicate two-tailed $P < .05$, and $P < .01$, respectively. ns: non significant.

Fig. 6: Impact of the histological subtype on textural features in non-small cell lung cancer patients

FLAB-derived textural features according to the histological subtypes (adenocarcinoma: ADC; squamous cell carcinoma: SqCC; large cell carcinoma: LCC) in non-small cell lung cancer patients for the three reconstructions used. CH_{AUC}: area under the curve of the cumulative histogram; high intensity larger area emphasis (HILAE); ZP: zone percentage. Data is shown as Tukey boxplots (lines displaying median, 25th and 75th percentiles; cross represents the mean value).

Supplemental Fig. 1

Illustration of the in-house software used to define a 3D box around the tumor, aiming at enclosing the tumor as well as excluding any nearby physiological or undesired uptake. For details see the materials and methods section.

Supplemental Fig. 2

Relationship between quantitative values extracted from PSF or PSF_7 and OSEM images, assessed using Bland-Altman plots for SUV_{max} (**a**) and SUV_{mean} (**b**) in tumour lesions.

Supplemental Fig. 3

Plots of TF features (CHAUC: area under the curve of the cumulative histogram; high intensity larger area emphasis (HILAE); ZP: zone percentage) against tumours MATV for OSEM versus PSF reconstructions.

Supplemental Fig. 4

Plots of TF features (CHAUC: area under the curve of the cumulative histogram; high intensity larger area emphasis (HILAE); ZP: zone percentage) against tumours MATV for OSEM versus PSF_7 reconstructions.

Supplemental Fig. 5: Impact of the EARL harmonization strategy on textural features using the FLAB algorithm independently on the three sets of reconstructions to delineate lesions

Textural features for the three reconstructions used. CH_{AUC}: area under the curve of the cumulative histogram; high intensity larger area emphasis (HILAE); ZP: zone percentage. Data is shown as Tukey boxplots (lines displaying median, 25th and 75th percentiles; cross represents the mean value).*, **, and *** indicate two-tailed $P < .05$, $P < .01$, and $P < .001$, respectively. ns : non significant. Data represent the 8 outliers (above the 90th percentiles) observed with OSEM-PSF MATV and/or PSF_7-PSF MATV when using a 50% of SUV_{max} threshold.

Supplemental Fig. 6

PET transverse slice of a PSF-reconstructed image showing one of the largest differences in MATV (-95%) between FLAB and 50% SUV_{max} thresholding.
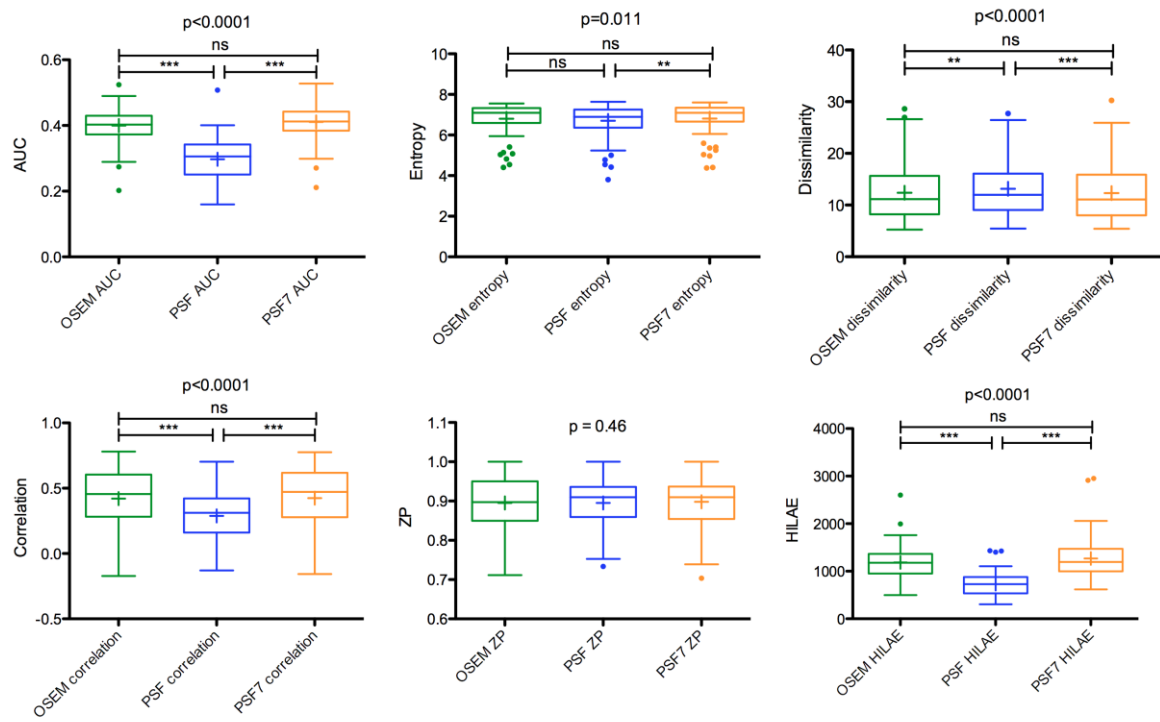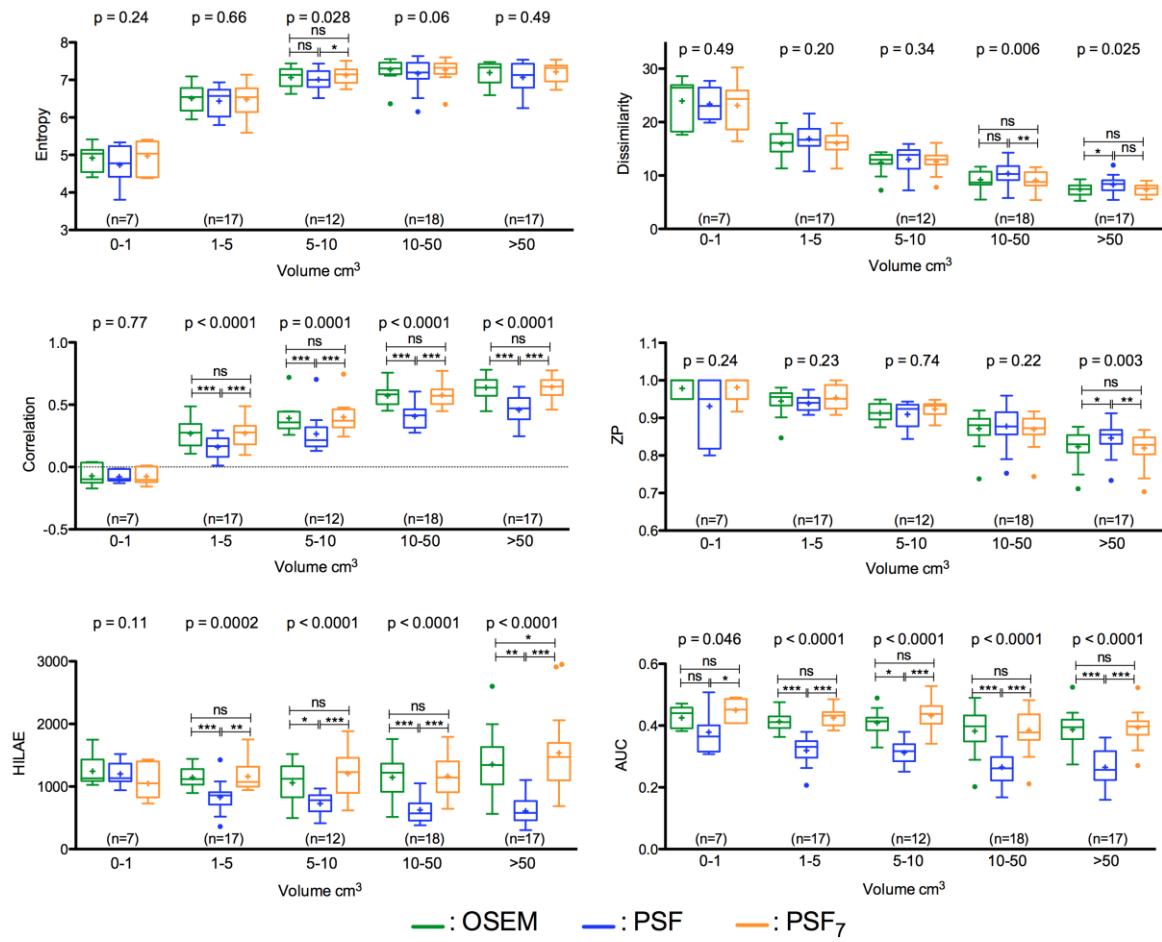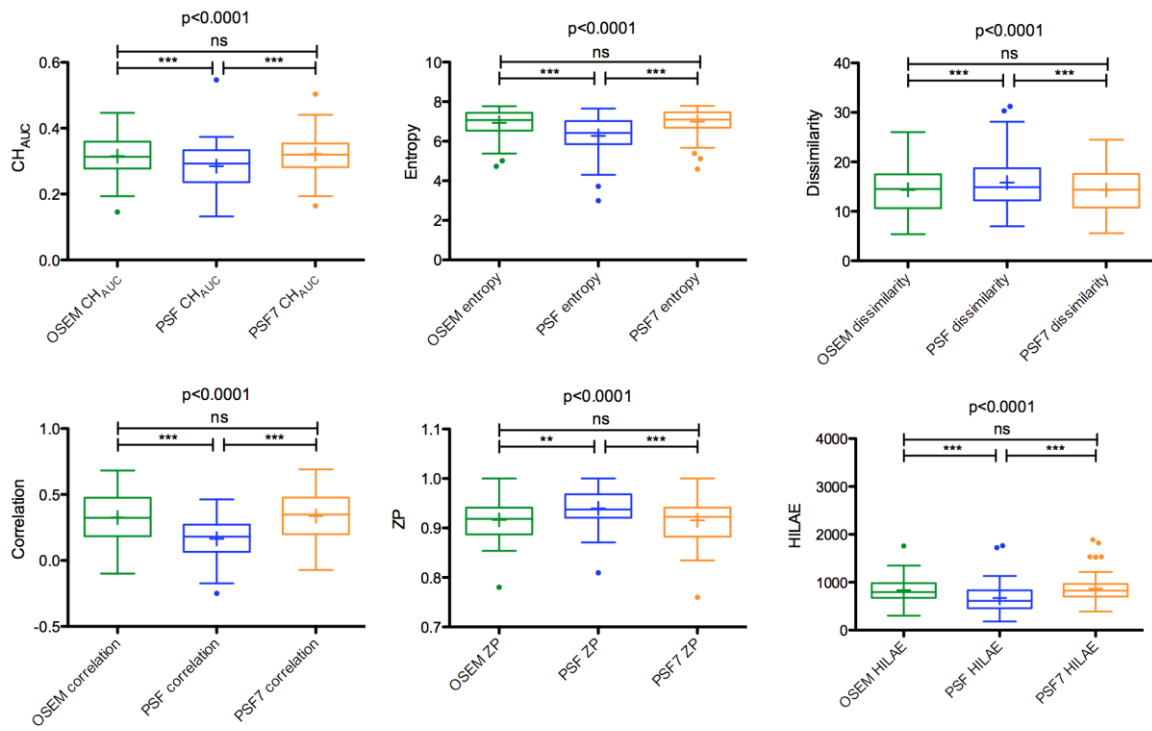
Fig. 1

Fig. 2

Fig. 3

Fig. 4

a

Metric	OSEM	PSF₇	PSF
SUV$_{max}$	4.5	4.6	5.5
SUV$_{mean}$	2.7	2.8	3.0
CH$_{AUC}$	0.36	0.35	0.34
Correlation	0.60	0.59	0.46
Dissimilarity	8.3	8.0	10.5
Entropy	7.3	7.2	7.5
HILAE	1009	924	880
ZP	0.87	0.87	0.88

MATV: 23.3 cm³

b

Metric	OSEM	PSF₇	PSF
MATV (cm³)	19.1	18.3	6.7
SUV$_{max}$	4.5	4.6	5.5
SUV$_{mean}$	2.9	3.0	4.1
CH$_{AUC}$	0.26	0.25	0.21
Correlation	0.49	0.47	0.27
Dissimilarity	10.2	9.7	11.1
Entropy	7.3	7.2	6.7
HILAE	612	564	360
ZP	0.88	0.90	0.93

c

Metric	OSEM	PSF₇	PSF
SUV$_{max}$	6.8	7.3	11.2
SUV$_{mean}$	4.3	4.5	5.2
CH$_{AUC}$	0.49	0.45	0.28
Correlation	0.50	0.51	0.31
Dissimilarity	10.8	11.6	11.1
Entropy	7.5	7.6	7.2
HILAE	1758	1372	687
ZP	0.89	0.92	0.87

MATV: 16.3 cm³

d

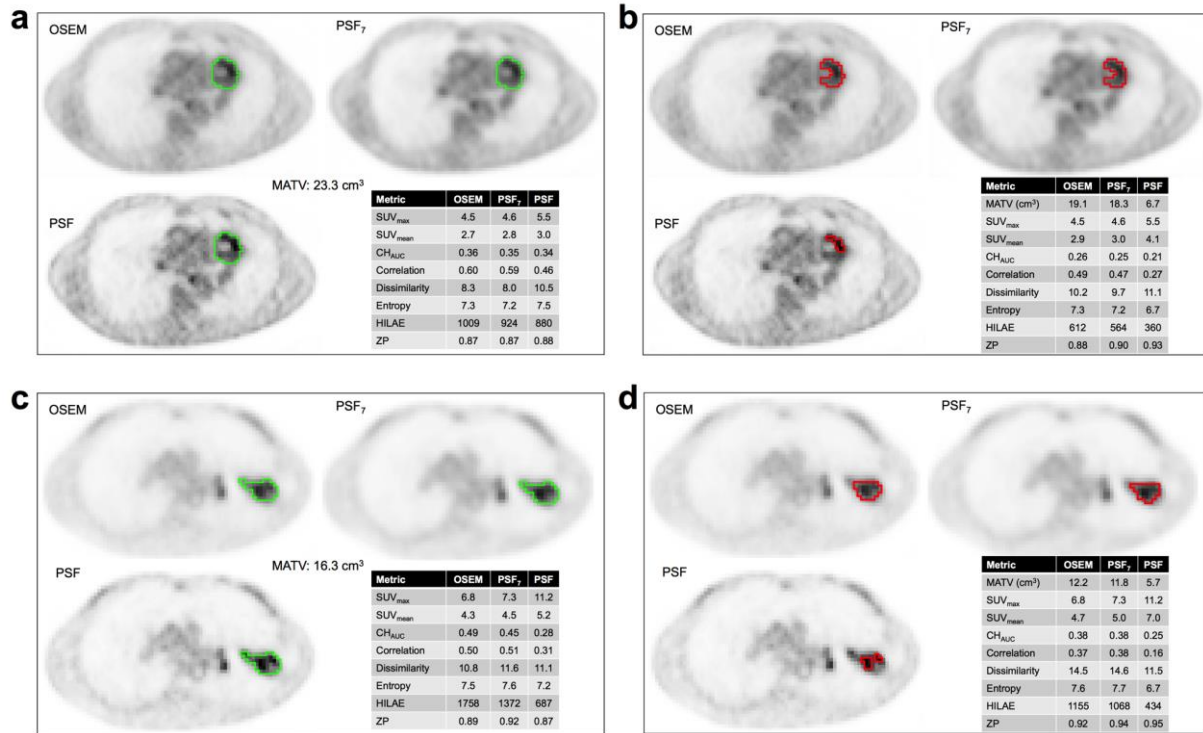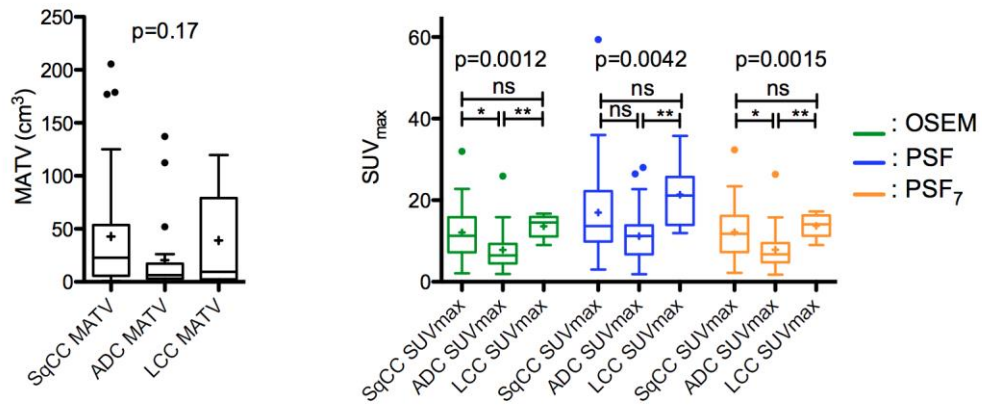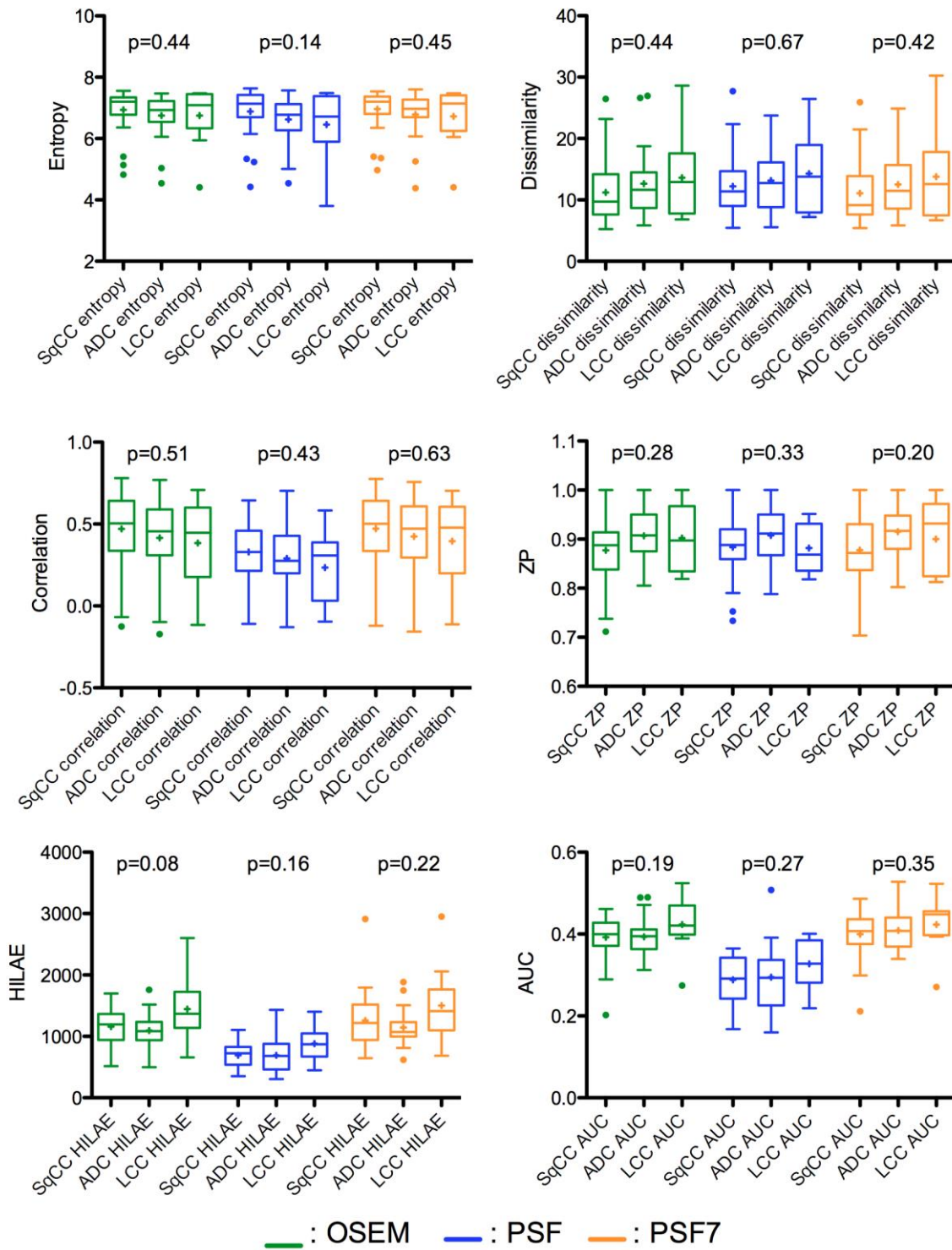| Metric	OSEM	PSF₇	PSF
MATV (cm³)	12.2	11.8	5.7
SUV$_{max}$	6.8	7.3	11.2
SUV$_{mean}$	4.7	5.0	7.0
CH$_{AUC}$	0.38	0.38	0.25
Correlation	0.37	0.38	0.16
Dissimilarity	14.5	14.6	11.5
Entropy	7.6	7.7	6.7
HILAE	1155	1068	434
ZP	0.92	0.94	0.95

Fig. 5

Fig. 6

References

1.   Boellaard R. Standards for PET image acquisition and quantitative data analysis. J Nucl Med. 2009;50 Suppl 1:11S-20S. doi:10.2967/jnumed.108.057182.

2.   Boellaard R. Methodological aspects of multicenter studies with quantitative PET. Methods Mol Biol. 2011;727:335-49. doi:10.1007/978-1-61779-062-1_18.

3.   Boellaard R. Mutatis mutandis: harmonize the standard! J Nucl Med. 2012;53:1-3. doi:10.2967/jnumed.111.094763.

4.   American College of Radiology, Nuclear Medicine and PET accreditation program http://www.acraccreditation.org/Modalities/Nuclear-Medicine-and-PET.

5.   European Association of Nuclear Medicine. EARL FDG-PET/CT accreditation. 2015. http://earl.eanm.org/cms/website.php?id=/en/projects/fdg_pet_ct_accreditation.htm.

6.   The Society of Nuclear Medicine and Molecular Imaging Clinical trials network (SNMMI-CTN). Scanner validation program. http://www.snmmi.org/Research/Content.aspx?ItemNumber=5482&navItemNumber=6834.

7.   Makris NE, Huisman MC, Kinahan PE, Lammertsma AA, Boellaard R. Evaluation of strategies towards harmonization of FDG PET/CT studies in multicentre trials: comparison of scanner validation phantoms and data analysis procedures. Eur J Nucl Med Mol Imaging. 2013;40:1507-15. doi:10.1007/s00259-013-2465-0.

8.   Panin VY, Kehren F, Michel C, Casey M. Fully 3-D PET reconstruction with system matrix derived from point source measurements. IEEE Trans Med Imaging. 2006;25:907-21.

9.   Teoh EJ, McGowan DR, Macpherson RE, Bradley KM, Gleeson FV. Phantom and clinical evaluation of the Bayesian penalized likelihood reconstruction algorithm Q.Clear on an LYSO PET/CT system. J Nucl Med. 2015. doi:10.2967/jnumed.115.159301.

10.   Bellevre D, Blanc Fournier C, Switsers O, Dugue AE, Levy C, Allouache D, et al. Staging the axilla in breast cancer patients with (1)(8)F-FDG PET: how small are the metastases that we can detect with new generation clinical PET systems? Eur J Nucl Med Mol Imaging. 2014;41:1103-12. doi:10.1007/s00259-014-2689-7.

11.   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;42:328-54. doi:10.1007/s00259-014-2961-x.

12.   Lasnon C, Desmonts C, Quak E, Gervais R, Do P, Dubos-Arvis C, et al. Harmonizing SUVs in multicentre trials when using different generation PET systems: prospective validation in non-small cell lung cancer patients. Eur J Nucl Med Mol Imaging. 2013;40:985-96. doi:10.1007/s00259-013-2391-1.

13.   Quak E, Le Roux PY, Hofman MS, Robin P, Bourhis D, Callahan J, et al. Harmonizing FDG PET quantification while maintaining optimal lesion detection: prospective multicentre validation in 517 oncology patients. Eur J Nucl Med Mol Imaging. 2015. doi:10.1007/s00259-015-3128-0.

14.   Mertens J, Dobbeleir A, Ham H, D'Asseler Y, Goethals I, Van de Wiele C. Standardized added metabolic activity (SAM): a partial volume independent marker of total lesion glycolysis in liver metastases. Eur J Nucl Med Mol Imaging. 2012;39:1441-8. doi:10.1007/s00259-012-2166-0.

15.   Visvikis D, Hatt M, Tixier F, Cheze Le Rest C. The age of reason for FDG PET image-derived indices. Eur J Nucl Med Mol Imaging. 2012;39:1670-2. doi:10.1007/s00259-012-2239-0.

16.   El Naqa I, Grigsby P, Apte A, Kidd E, Donnelly E, Khullar D, et al. Exploring feature-based approaches in PET images for predicting cancer treatment outcomes. Pattern Recognit. 2009;42:1162-71. doi:10.1016/j.patcog.2008.08.011.

17.   Galavis PE, Hollensen C, Jallow N, Paliwal B, Jeraj R. Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters. Acta Oncol. 2010;49:1012-6. doi:10.3109/0284186X.2010.498437.

18.   Yan J, Chu-Shern JL, Loi HY, Khor LK, Sinha AK, Quek ST, et al. Impact of Image Reconstruction Settings on Texture Features in 18F-FDG PET. J Nucl Med. 2015;56:1667-73. doi:10.2967/jnumed.115.156927.

19.   Hatt M, Majdoub M, Vallieres M, Tixier F, Le Rest CC, Groheux D, et al. 18F-FDG PET uptake characterization through texture analysis: investigating the complementary nature of heterogeneity and functional tumor volume in a multi-cancer site patient cohort. J Nucl Med. 2015;56:38-44. doi:10.2967/jnumed.114.144055.

20.   Calais J, Dubray B, Nkhali L, Thureau S, Lemarignier C, Modzelewski R, et al. High FDG uptake areas on pre-radiotherapy PET/CT identify preferential sites of local relapse after chemoradiotherapy for locally advanced oesophageal cancer. Eur J Nucl Med Mol Imaging. 2015;42:858-67. doi:10.1007/s00259-015-3004-y.

21.   Sauter AW, Schwenzer N, Divine MR, Pichler BJ, Pfannenberg C. Image-derived biomarkers and multimodal imaging strategies for lung cancer management. Eur J Nucl Med Mol Imaging. 2015;42:634-43. doi:10.1007/s00259-014-2974-5.

22.   Takeuchi S, Khiewvan B, Fox PS, Swisher SG, Rohren EM, Bassett RL, Jr., et al. Impact of initial PET/CT staging in terms of clinical stage, management plan, and prognosis in 592 patients with non-small-cell lung cancer. Eur J Nucl Med Mol Imaging. 2014;41:906-14. doi:10.1007/s00259-013-2672-8.

23.   Vera P, Mezzani-Saillard S, Edet-Sanson A, Menard JF, Modzelewski R, Thureau S, et al. FDG PET during radiochemotherapy is predictive of outcome at 1 year in non-small-cell lung cancer patients: a prospective multicentre study (RTEP2). Eur J Nucl Med Mol Imaging. 2014;41:1057-65. doi:10.1007/s00259-014-2687-9.

24.   Fried DV, Mawlawi O, Zhang L, Fave X, Zhou S, Ibbott G, et al. Stage III Non-Small Cell Lung Cancer: Prognostic Value of FDG PET Quantitative Imaging Features Combined with Clinical Prognostic Factors. Radiology. 2016;278:214-22. doi:10.1148/radiol.2015142920.

25.   Lovinfosse P, Janvary ZL, Coucke P, Jodogne S, Bernard C, Hatt M, et al. FDG PET/CT texture analysis for predicting the outcome of lung cancer treated by stereotactic body radiation therapy. Eur J Nucl Med Mol Imaging. 2016. doi:10.1007/s00259-016-3314-8.

26.   Aide N, Desmonts C, Beauregard JM, Beyer T, Kinross K, Roselt P, et al. High throughput static and dynamic small animal imaging using clinical PET/CT: potential preclinical applications. Eur J Nucl Med Mol Imaging. 2010;37:991-1001. doi:10.1007/s00259-009-1352-1.

27.   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. doi:10.1109/TMI.2008.2012036.

28.   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;38:663-72. doi:10.1007/s00259-010-1688-6.

29.   Hatt M, Cheze-Le Rest C, Aboagye EO, Kenny LM, Rosso L, Turkheimer FE, et al. Reproducibility of 18F-FDG and 3'-deoxy-3'-18F-fluorothymidine PET tumor volume measurements. J Nucl Med. 2010;51:1368-76. doi:10.2967/jnumed.110.078501.

30.   Hatt M, Cheze le Rest C, Descourt P, Dekker A, De Ruysscher D, Oellers M, et al. Accurate automatic delineation of heterogeneous functional volumes in positron emission tomography for oncology applications. Int J Radiat Oncol Biol Phys. 2010;77:301-8. doi:10.1016/j.ijrobp.2009.08.018.

31.   Arens AI, Troost EG, Hoeben BA, Grootjans W, Lee JA, Gregoire V, 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. doi:10.1007/s00259-013-2651-0.

32.   van Velden FH, Cheebsumon P, Yaqub M, Smit EF, Hoekstra OS, Lammertsma AA, et al. Evaluation of a cumulative SUV-volume histogram method for parameterizing heterogeneous intratumoural FDG uptake in non-small cell lung cancer PET studies. Eur J Nucl Med Mol Imaging. 2011;38:1636-47. doi:10.1007/s00259-011-1845-6.

33.   Hatt M, Tixier F, Cheze Le Rest C, Pradier O, Visvikis D. Robustness of intratumour (1)(8)F-FDG PET uptake heterogeneity quantification for therapy response prediction in oesophageal carcinoma. Eur J Nucl Med Mol Imaging. 2013;40:1662-71. doi:10.1007/s00259-013-2486-8.

34.   Tixier F, Hatt M, Le Rest CC, Le Pogam A, Corcos L, Visvikis D. Reproducibility of tumor uptake heterogeneity characterization through textural feature analysis in 18F-FDG PET. J Nucl Med. 2012;53:693-700. doi:10.2967/jnumed.111.099127.

35.   Vallieres M, Freeman CR, Skamene SR, El Naqa I. A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. Phys Med Biol. 2015;60:5471-96. doi:10.1088/0031-9155/60/14/5471.

36.   Webb AG. Introduction to Biomedical Imaging. : Wiley-IEEE Press; 2003.

37.   Nyflot MJ, Yang F, Byrd D, Bowen SR, Sandison GA, Kinahan PE. Quantitative radiomics: impact of stochastic effects on textural feature analysis implies the need for standards. J Med Imaging (Bellingham). 2015;2:041002. doi:10.1117/1.JMI.2.4.041002.

38.   Sunderland JJ, Christian PE. Quantitative PET/CT scanner performance characterization based upon the society of nuclear medicine and molecular imaging clinical trials network oncology clinical simulator phantom. J Nucl Med. 2015;56:145-52. doi:10.2967/jnumed.114.148056.

39.   van Velden FH, Kramer GM, Frings V, Nissen IA, Mulder ER, de Langen AJ, et al. Repeatability of Radiomic Features in Non-Small-Cell Lung Cancer [F]FDG-PET/CT Studies: Impact of Reconstruction and Delineation. Molecular imaging and biology : MIB : the official publication of the Academy of Molecular Imaging. 2016. doi:10.1007/s11307-016-0940-2.

40.   Groheux D, Majdoub M, Tixier F, Le Rest CC, Martineau A, Merlet P, et al. Do clinical, histological or immunohistochemical primary tumour characteristics translate into different (18)F-FDG PET/CT volumetric and heterogeneity features in stage II/III breast cancer? Eur J Nucl Med Mol Imaging. 2015;42:1682-91. doi:10.1007/s00259-015-3110-x.

41.   Ha S, Choi H, Cheon GJ, Kang KW, Chung JK, Kim EE, et al. Autoclustering of Non-small Cell Lung Carcinoma Subtypes on (18)F-FDG PET Using Texture Analysis: A Preliminary Result. Nucl Med Mol Imaging. 2014;48:278-86. doi:10.1007/s13139-014-0283-3.

42.   Chung C. Tyrosine kinase inhibitors for epidermal growth factor receptor gene mutation-positive non-small cell lung cancers: an update for recent advances in therapeutics. J Oncol Pharm Pract. 2015. doi:10.1177/1078155215577810.

43.   Caicedo C, Garcia-Velloso MJ, Lozano MD, Labiano T, Vigil Diaz C, Lopez-Picazo JM, et al. Role of [(1)(8)F]FDG PET in prediction of KRAS and EGFR mutation status in patients with advanced non-small-cell lung cancer. Eur J Nucl Med Mol Imaging. 2014;41:2058-65. doi:10.1007/s00259-014-2833-4.

44.   Ko KH, Hsu HH, Huang TW, Gao HW, Shen DH, Chang WC, et al. Value of (1)(8)F-FDG uptake on PET/CT and CEA level to predict epidermal growth factor receptor mutations in pulmonary adenocarcinoma. Eur J Nucl Med Mol Imaging. 2014;41:1889-97. doi:10.1007/s00259-014-2802-y.