



Implications of reconstruction protocol for histo-biological characterisation of breast cancers using FDG-PET radiomics

Nicolas Aide, Thibault Salomon, Cécile Blanc-Fournier, Jean-Michel Grellard,
Christelle Levy, Charline Lasnon

► To cite this version:

Nicolas Aide, Thibault Salomon, Cécile Blanc-Fournier, Jean-Michel Grellard, Christelle Levy, et al.. Implications of reconstruction protocol for histo-biological characterisation of breast cancers using FDG-PET radiomics. EJNMMI Research, 2018, 8 (1), pp.114. 10.1186/s13550-018-0466-5 . inserm-02108605

HAL Id: inserm-02108605

<https://inserm.hal.science/inserm-02108605>

Submitted on 24 Apr 2019

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.

ORIGINAL RESEARCH

Open Access



Implications of reconstruction protocol for histo-biological characterisation of breast cancers using FDG-PET radiomics

Nicolas Aide^{1,2†}, Thibault Salomon^{1†}, Cécile Blanc-Fournier³, Jean-Michel Grellard⁴, Christelle Levy⁵ and Charline Lasnon^{2,6*} 

Abstract

Background: The aim of this study is to determine if the choice of the ^{18}F -FDG-PET protocol, especially matrix size and reconstruction algorithm, is of importance to discriminate between immunohistochemical subtypes (luminal versus non-luminal) in breast cancer with textural features (TFs).

Procedures: Forty-seven patients referred for breast cancer staging in the framework of a prospective study were reviewed as part of an ancillary study. In addition to standard PET imaging ($\text{PSF}_{\text{WholeBody}}$), a high-resolution breast acquisition was performed and reconstructed with OSEM and PSF ($\text{OSEM}_{\text{breast}}/\text{PSF}_{\text{breast}}$). PET standard metrics and TFs were extracted. For each reconstruction protocol, a prediction model for tumour classification was built using a random forests method. Spearman coefficients were used to seek correlation between PET metrics.

Results: $\text{PSF}_{\text{WholeBody}}$ showed lower numbers of voxels within VOIs than $\text{OSEM}_{\text{breast}}$ and $\text{PSF}_{\text{breast}}$ with median (interquartile range) equal to 130 (43–271), 316 (167–1042), 367 (107–1221), respectively ($p < 0.0001$). Therefore, using LifeX software, 28 (59%), 46 (98%) and 42 (89%) patients were exploitable with $\text{PSF}_{\text{WholeBody}}$, $\text{OSEM}_{\text{breast}}$ and $\text{PSF}_{\text{breast}}$, respectively.

On matched comparisons, $\text{PSF}_{\text{breast}}$ reconstruction presented better abilities than $\text{PSF}_{\text{WholeBody}}$ and $\text{OSEM}_{\text{breast}}$ for the classification of luminal versus non-luminal breast tumours with an accuracy reaching 85.7% as compared to 67.8% for $\text{PSF}_{\text{WholeBody}}$ and 73.8% for $\text{OSEM}_{\text{breast}}$. $\text{PSF}_{\text{breast}}$ accuracy, sensitivity, specificity, PPV and NPV were equal to 85.7%, 94.3%, 42.9%, 89.2%, 60.0%, respectively. Coarseness and ZLNU were found to be main variables of importance, appearing in all three prediction models. Coarseness was correlated with SUV_{max} on $\text{PSF}_{\text{WholeBody}}$ images ($p = -0.526$, $p = 0.005$), whereas it was not on $\text{OSEM}_{\text{breast}}$ ($p = -0.183$, $p = 0.244$) and $\text{PSF}_{\text{breast}}$ ($p = -0.244$, $p = 0.119$) images. Moreover, the range of its values was higher on $\text{PSF}_{\text{breast}}$ images as compared to $\text{OSEM}_{\text{breast}}$, especially in small lesions ($\text{MTV} < 3 \text{ ml}$).

Conclusions: High-resolution breast PET acquisitions, applying both small-voxel matrix and PSF modelling, appeared to improve the characterisation of breast tumours.

Keywords: ^{18}F -FDG-PET, Breast cancer, Heterogeneity, Radiomics, Reconstruction

* Correspondence: clasnon@baclesse.unicancer.fr

[†]Nicolas Aide and Thibault Salomon contributed equally to this work.

²INSERM 1199 ANTICIPE, Normandy University, Caen, France

⁶Nuclear Medicine Department, François Baclesse Cancer Centre, 3 Avenue du Général Harris, BP 45026 Cedex 5, 14076 Caen, France

Full list of author information is available at the end of the article

Background

Breast cancer is the most common type of cancer and the leading cause of death related to cancer in women worldwide [1]. It displays a large inter- and intra-tumour heterogeneity with a strong impact on patient management and outcome. Inter-patient tumoral heterogeneity can be reflected by actual staging systems and histopathological classifications that are predictors of patients' outcomes and major determinants for treatment planning [2]. In the context of invasive breast cancer staging, 2-deoxy-2-[^{18}F]-fluoro-D-glucose (^{18}F -FDG) positron emission tomography coupled with computed tomography (PET/CT) has shown its value for the detection of unexpected node involvements and/or distant metastasis [3]. Therefore, the European Society for Medical Oncology (ESMO) international consensus as well as the National Comprehensive Cancer Network (NCCN) guidelines recommend to consider the use of FDG PET-CT if available, instead of CT and bone scan for the initial staging of inoperable and non-metastatic locally advanced breast cancer (stage III with the exception of T3 N1) [4]. However, due to the high heterogeneity of breast cancers, FDG tumour uptake intensity measured as maximum standardised uptake value (SUV_{max}) is highly variable, depending on multiple factors such as histological type, phenotypic type [5], proliferation index [6], histological grade and the presence of a P53 mutation [7] for example. However, SUV_{max} has been shown to be a prognostic index in invasive breast cancer [5]. More recently, PET textural features have emerged in the field of cancerology and have shown promising results in predicted response to treatment and/or patient survival in cervix, head and neck, lung and oesophageal cancer [8–16]. In breast cancers, heterogeneous tumour FDG uptake appeared to be frequent, especially in large tumours with intense FDG uptake [17]. Some studies have demonstrated that FDG breast tumour heterogeneity, based on single parameter or multi-feature signature, is significantly correlated with immunohistochemical factors and St Gallen's subtypes [18–20]. Interestingly, these heterogeneity parameters were not correlated to SUV, meaning that they can surely provide additional information. However, these results are controversial because other studies did not find any ability of textural features (TFs) to discriminate between immunohistochemical subtypes [21, 22]. It is worth noticing that neither of these two studies used dedicated high-resolution images but images with standard $4 \times 4 \times 4$ mm voxels. These findings thus suggest a potential role of textural features in breast cancer for non-invasive molecular subtype classification and subsequent patient prognosis stratification, but PET procedure seems to arise as a critical point in this field, especially when considering breast tumours that are usually small. Indeed, for such small lesions, it had already been demonstrated that small-voxel reconstruction and latest reconstruction algorithms bring better signal-

to-noise ratio and could improve tumoral detection and the sensitivity of visual lymph node characterisation [23–25]. Therefore, the aim of this ancillary prospective clinical study is to compare different PET protocols with regard to their ability to discriminate between luminal versus non-luminal breast tumours.

Material and methods

Study population

This study is ancillary to a previous monocentric prospective study conducted by our team and approved by the local Ethics Committee (CPP Nord Ouest III, reference 2009-10) [23]. Informed and signed consent was obtained from all patients. Patients with newly diagnosed and histologically proven breast cancer for which breast surgery and axillary lymph node dissection was indicated were included from April 2009 to June 2012. All patients had a ^{18}F -FDG PET/CT for initial staging of the disease.

PET/CT acquisitions

PET imaging studies were performed on a Biograph TrueV (Siemens Medical Solutions). ^{18}F -FDG injection was preceded by a 6-h fasting period and a 15-min rest in a warm room. Patients were scanned 60 min after ^{18}F -FDG injection from the skull base to the mid-thighs (2 min 40 s per bed position for normal-weight patients ($\text{BMI} \leq 25 \text{ kg/m}^2$) and 3 min and 40 s per bed position for patients with $\text{BMI} > 25 \text{ kg/m}^2$). Images were reconstructed using a point spread function (PSF) algorithm (HD; TrueX, Siemens Medical Solutions, 3 iterations and 21 subsets) with no post-filtering ($\text{PSF}_{\text{WholeBody}}$) and a 168^2 matrix size leading to a voxel size of $4.1 \times 4.1 \times 5.0$ mm. A complementary high-resolution (HR) breast dedicated bed position (6 min per bed position) was performed just after the completion of the skull base to the mid-thighs acquisition. Images were reconstructed using the same protocol as above ($\text{PSF}_{\text{breast}}$) and an ordered subset expectation maximization (OSEM) algorithm (four iterations, eight subsets) with a Gaussian post-filtering of 5 mm ($\text{OSEM}_{\text{breast}}$) with a 512^2 matrix size leading to voxels of $1.3 \times 1.3 \times 1.9$ mm. Scatter and attenuation corrections were applied for both acquisitions.

PET/CT images analysis

Injected dose, time between injection and acquisition and capillary glycaemia were recorded to seek for EANM recommendations fulfilment [26]. A single nuclear medicine physician drew volumes of interest (VOIs) encompassing the entire breast tumour on each PET acquisition using a PET edge method implemented in MIM software (MIM software, Cleveland, OH, US, version 5.6.5). In case of multiple lesions, only the biggest lesion was considered. To be close to real clinical

practice, each PET dataset was contoured independently as it would have been done in a PET unit. The PET gradient method was used because it had been shown to be reproducible, little impacted by reconstruction type and have the ability to encompass the entire tumour by taking into account cold zones as opposed to threshold based VOIs [27]. Moreover, it is widely available. VOIs were subsequently saved as DICOM RT structures and then loaded in LifeX software [28] to extract SUV_{max} , SUV_{mean} , metabolic tumour volumes (MTV), total lesion glycolysis (TLG) and TFs parameters.

The following TFs were extracted:

- Homogeneity, energy, contrast, correlation, entropy, dissimilarity from grey level co-occurrence matrix (GLCM) that takes into account the arrangements of pairs of voxels
- Coarseness, contrast and busyness from neighbourhood grey-level different matrix (NGLDM) that corresponds to the difference of grey-level between one voxel and its 26 neighbours in 3 dimensions.
- SZE, LZE, LGZE, HGZE, SZLGE, SZHGE, LZLGE, LZHGE, GLNU, ZLNU, ZP (Table 1) from grey-level zone length matrix (GLZLM) that provides information on the size of homogeneous zones for each grey-level in three dimensions.

Absolute resampling using 64 grey levels between 0 and the maximum SUV units recorded for each reconstruction was used for all TFs: 27 for $PSF_{WholeBody}$, 15 for $OSEM_{breast}$ and 32 for PSF_{breast} leading to a size of bin of 0.4, 0.2 and 0.5, respectively [29, 30].

Coefficients of variation (CoV), measured in a 4 cm³ spherical VOI set in the descending thoracic aorta, were computed as follows for each reconstruction protocol:

$$CoV = \frac{\text{standard deviation}}{SUV_{\text{mean}}}$$

Further analyses were undergone. First, to assess the impact of quantification scaling, a supplemental analysis was undergone by using an upper SUV bound set to 32 for all 3 reconstructions leading to a size of bin of 0.5 for all reconstructions. Secondly, to assess the impact of the voxel size, a post-reconstruction resampling was applied to $PSF_{wholeBody}$ and to PSF_{breast} images to obtain a 2 mm³ voxel size and a 4 mm³ voxel size, respectively.

Statistical analysis

Quantitative data are presented as the median (interquartile range) or the mean (SD) when appropriate.

To compare PET metrics extracted from the three different reconstructions, non-parametric Friedman test with post-hoc test were used.

For each reconstruction protocol, a random forests (RF) method was used for building a prediction model for luminal versus non-luminal tumour classification. The method implemented classification and regression trees (CART, $n = 100$) and bootstrapping aggregating (bagging) method proposed by Breiman [31–33]. It allows studying the global heterogeneity of tumour rather than looking at individual features. For the validation, i.e. the training accuracy, the internal check in RF itself was used, based on the prediction error using the Out-Of-Bag (OOB) estimates of classification error: the smaller the OOB error rate, the better the reconstruction is able to classify between luminal and non-luminal tumours. Sensitivity (Se), specificity (Sp), positive predictive value (PPV), negative predictive value (NPV) and accuracy were computed. The importance of TFs in classification was assessed for each reconstruction protocol by measuring the mean decrease accuracy [34] of class prediction. Spearman coefficients were used to seek correlation between PET metrics of importance. Finally, the first three main PET metrics were considered for further paired comparison between reconstruction protocols using Friedman test with post-hoc test, Spearman correlation tests and ROC analyses.

Graph and statistical analysis were performed on XLSTAT Software (XLSTAT 2017: Data Analysis and Statistical Solution for Microsoft Excel. Addinsoft, Paris, France (2017)). For all statistical tests, a two-tailed P value of less than 0.05 was considered statistically significant.

Results

Patients and PET characteristics

Sixty-three patients were referred for the staging of breast carcinoma from April 2009 to June 2012. Sixteen patients were excluded from the analysis, for a final

Table 1 Third-order textural features abbreviations

Variables	Definition
SZE	Short-zone emphasis
LZE	Long-zone emphasis
LGZE	Low grey-level zone emphasis
HGZE	High grey-level zone emphasis
SZLGE	Short-zone low grey-level zone emphasis
SZHGE	Short-zone high grey-level zone emphasis
LZLGE	Long-zone low grey-level zone emphasis
LZHGE	Long-zone high grey-level zone emphasis
GLNU	Grey-level non-uniformity
ZLNU	Zone length non-uniformity
ZP	Zone percentage

database of 47 patients (47 PET-CTs). The causes of exclusion were as follows: PET-CT not performed prior to surgery ($n=8$), metastatic tumours on staging imaging ($n=4$), missing data ($n=1$) and breast lesions not visible on PET-CT ($n=3$). The tumour types confirmed on histopathology included 38 infiltrating ductal carcinomas, 2 infiltrating lobular carcinomas, 5 mixed ductal/lobular infiltrating carcinomas, 1 tubular carcinoma and 1 infiltrating undifferentiated carcinoma. Patient characteristics are displayed in Table 2. Six patients (12.8%) had multiple lesions (range 2–4).

The mean injected dose was equal to 282.1 (67.5) MBq. The mean uptake time was 62.0 (7.7) min for the whole body and 81.9 (8.3) min for HR breast acquisitions.

Lesions size: metabolic tumour volumes and number of voxels within the VOIs

Median metabolic tumour volumes (MTVs) ($n=47$) were 3.62 (1.29–9.85), 2.43 (1.14–8.68), 2.43 (0.78–10.17) ml for PSF_{WholeBody}, OSEM_{breast} and PSF_{breast}, respectively ($p<0.00001$). Thirty five patients (74.5%) had MTVs $<10\text{ cm}^3$ with at least one PET protocol. Dedicated HR breast acquisitions led to significantly smaller MTVs than PSF_{WholeBody} acquisitions for both OSEM_{breast} ($p=0.037$) and PSF_{breast} ($p<0.0001$). There was no significant difference between PSF_{breast} and OSEM_{breast} MTVs ($p=0.079$) (Fig. 1a). The median numbers of voxels within VOIs were 130 (43–271), 316 (167–1042), 367 (107–1221) for PSF_{WholeBody}, OSEM_{breast} and PSF_{breast}, respectively ($p<0.0001$). Dedicated HR breast acquisitions led to a significantly higher number of voxels than PSF_{WholeBody} acquisitions for both OSEM_{breast} ($p<0.0001$) and PSF_{breast} ($p<0.0001$) reconstructions. There was no significant difference between PSF_{breast} and OSEM_{breast} numbers of voxels ($p=0.062$) (Fig. 1b). To be analysed in LifeX software, MTVs should contain at least 64 voxels. Therefore, 28 (59%), 46 (98%) and 42 (89%) patients were exploitable when using PSF_{WholeBody}, OSEM_{breast} and PSF_{breast} reconstructions, respectively. Of note, due to a very low MTV, only one patient was not analysable by all three reconstructions and therefore was not included in the subsequent statistical analysis. She was a 72-year-old woman presenting a luminal A (ER+; PR+, HER2–, grade 1) breast tumour classified T1N1M0.

Prediction accuracies for luminal status tumours classification and variables of importance identification for PSF_{WholeBody}, PSF_{breast} and OSEM_{breast} PET protocols

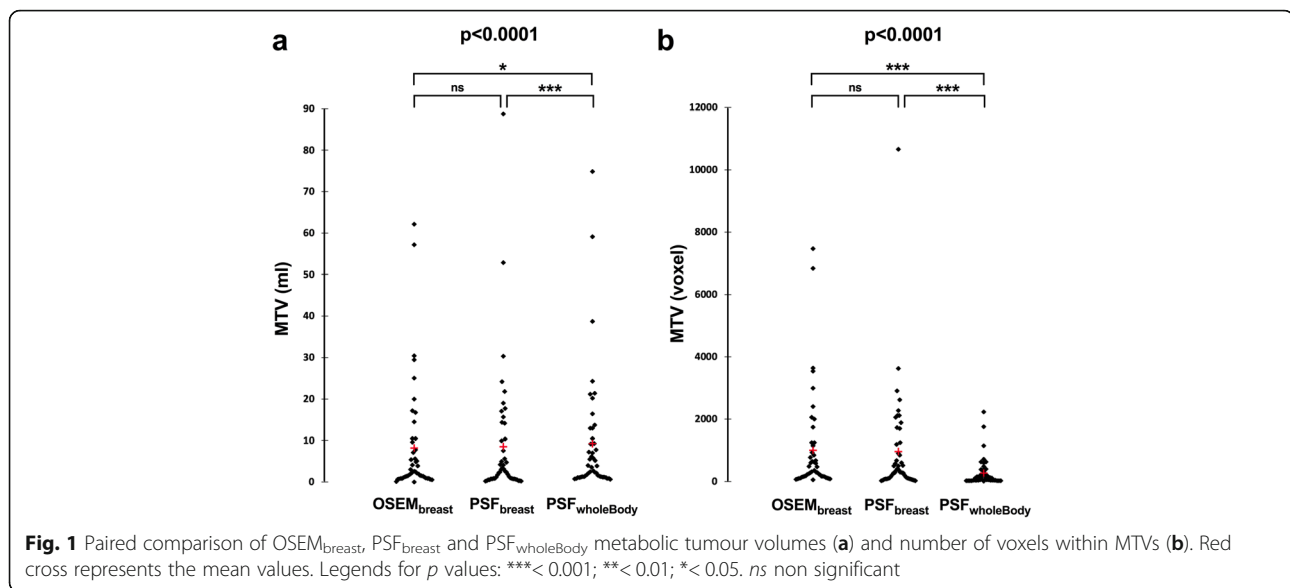
When matched comparing the 28 patients analysed with PSF_{WholeBody} and PSF_{breast} (22 luminal and 6 non-luminal tumours), PSF_{WholeBody} showed higher OOB estimates of classification error than PSF_{breast} with values equal to 32.1% and 25.0%, respectively.

Table 2 Patients' characteristics

Characteristics	Number of patients ($n=47$)
Age (years), mean \pm SD [min-max]	56.1 \pm 12.3 [29–80]
Histologic type, n (%)	
Ductal	38 (80.8)
Lobular	2 (4.3)
Mixed type	5 (10.6)
Others	2 (4.3)
Estrogen receptor status, n (%)	
Positive	40 (85.1)
Negative	7 (14.9)
Progesterone receptor status, n (%)	
Positive	33 (70.2)
Negative	14 (29.8)
HER2 status	
Positive	6 (12.8)
Negative	41 (87.2)
Mitotic grade	
1–2	29 (61.7)
3	18 (38.3)
T, n (%)	
T1	13 (27.6)
T2	24 (51.1)
T3	10 (21.3)
N, n (%)	
0	14 (29.8)
1	24 (51.1)
2	5 (10.6)
3	4 (8.5)
AJCC stage, n (%)	
I	4 (8.5)
IIA	16 (34.0)
IIB	11 (23.4)
IIIA	12 (25.6)
IIIB	0 (0)
IIIC	4 (8.5)

AJCC American Joint Committee on Cancer

Accuracy, Se, Sp, PPV and NPV are displayed in Table 3 and variables of importance for both PET protocols are displayed on Fig. 2. Interestingly, both protocols found coarseness and ZLNU to be variables of importance. However, coarseness was negatively correlated with SUV_{max} and SUV_{mean} on PSF_{WholeBody} images ($\rho=-0.526$, $p=0.005$ and $\rho=-0.406$, $p<0.033$, respectively), whereas it was not on PSF_{breast} images ($\rho=-0.093$, $p=0.636$ and $\rho=0.139$, $p<0.479$, respectively). ZLNU



was correlated with SUV_{max} and SUV_{mean} on both PSF_{wholeBody} and PSF_{breast} images ($\rho = 0.928$, $p < 0.0001$ and $\rho = 0.962$, $p < 0.0001$, respectively for SUV_{max} correlation). Noticeably, variables of importance mean decrease accuracies were globally lower for PSF_{wholeBody} images as compared to PSF_{breast} images. Moreover, all textural features were highly correlated with SUV_{max} and SUV_{mean} values for PSF_{wholeBody} images, whereas correlations were fewer and lower for PSF_{breast} images (Fig. 2). Concerning images noise, there was no significant difference between PSF_{wholeBody} and PSF_{breast} images with a mean CoV of 0.175 (0.030) and 0.189 (0.031), respectively ($p = 0.087$). Moreover, coarseness was not correlated to noise: $\rho = -0.029$, $p = 0.883$ for PSF_{wholeBody} and $\rho = 0.190$, $p = 0.330$ for PSF_{breast}.

When applying a size of bin equal to 0.5 on PSF_{wholeBody} images to meet the quantification scale of PSF_{breast}, the OOB estimates of classification error went from 32.1 to 28.6%, still higher than PSF_{breast} OOB estimates of classification error. Variables of importance and their correlations are displayed on Additional file 1: Figure S1a.

After applying to PSF_{wholeBody} images a 2 mm³ post-reconstruction voxel resampling, the OOB estimates of classification error remained stable, equal to 32.1%. Variables of importance and their correlations for both

protocols are displayed on Additional file 2: Figure S2a. Interestingly, there was no more correlation between SUV_{max} and coarseness values on PSF_{wholeBody} images after post-reconstruction voxel resampling: $\rho = -0.276$, $p = 0.155$.

When matched comparing the 42 patients analysed with both OSEM_{breast} and PSF_{breast} (35 luminal and 7 non-luminal tumours), PSF_{breast} showed higher classification accuracy and lower OOB estimates of classification error than OSEM_{breast}. OOB estimates were equal to 26.2% and 14.3% when using OSEM_{breast} and PSF_{breast}, respectively. Accuracy, Se, Sp, PPV and NPV are displayed in Table 3. Both protocols showed high sensitivity but low specificity for the luminal status detection: the best specificity was obtained using PSF_{breast} with a value equal to 42.9%. Figure 3 displays variables of importance for both PET protocols and demonstrates that coarseness and ZLNU were again variables of importance with both protocols as well as GLNU, SZLGE and busyness. GLNU and ZLNU were found to be significantly positively correlated with each other ($p < 0.0001$) and with SUV_{max} ($p < 0.0001$) on both protocols. Coarseness was not correlated with SUV_{max} with ρ equal to -0.183 ($p = 0.244$) for OSEM_{breast} and -0.244 ($p = 0.119$) for

Table 3 Reconstruction protocols classification performances regarding tumour luminal status detection

		Accuracy (%)	Se (%)	Sp (%)	PPV (%)	NPV (%)
Comparison of 28 patients	PSF _{wholeBody}	67.8	86.4	0.0	76.0	0.0
	PSF _{breast}	75.0	90.9	16.7	80.0	33.3
Comparison of 42 patients	OSEM _{breast}	73.8	85.7	14.3	83.3	16.7
	PSF _{breast}	85.7	94.3	42.9	89.2	60.0

Se sensitivity, Sp specificity, PPV positive predictive value, NPV negative predictive value

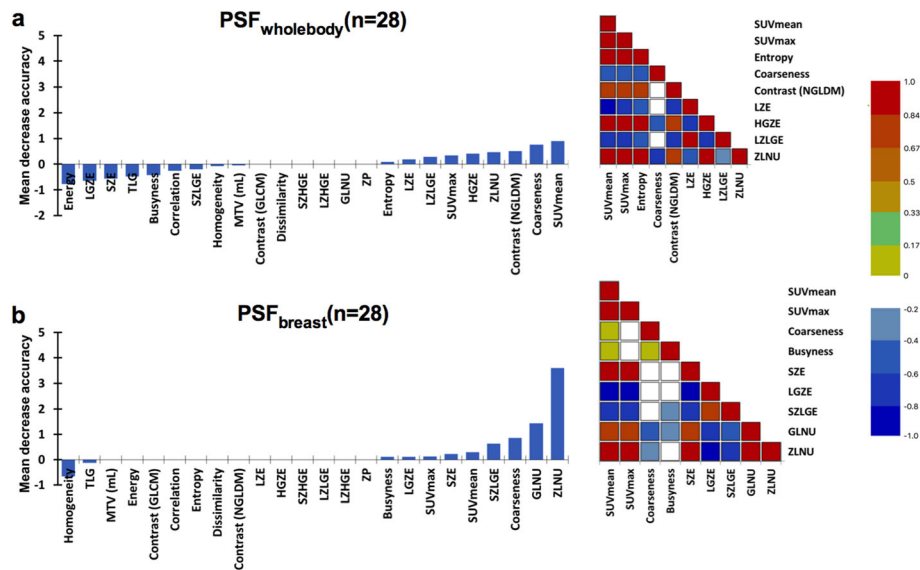


Fig. 2 Left panels display the mean decrease accuracy of textural features values and right panels display Spearman correlation matrixes of all PET metrics found to have positive mean decrease accuracy, whatever the value for PSF_{wholebody} (a) and PSF_{breast} (b) reconstructions. For Spearman correlation matrixes, the blue colour corresponds to a correlation close to -1 and the red colour corresponds to a correlation close to 1. The green corresponds to a correlation close to 0

PSF_{breast}. Concerning images noise, there was no significant difference between OSEM_{breast} and PSF_{breast} images with a mean CoV of 0.175 (0.030) and 0.189 (0.031), respectively ($p = 0.087$). Moreover, coarseness was not correlated to noise: $\rho = -0.029$, $p = 0.883$ for PSF_{wholebody} and $\rho = 0.190$, $p = 0.330$ for PSF_{breast}.

When applying a size of bin equal to 0.5 on OSEM_{breast} images to meet the quantification scale of PSF_{breast}, the OOB estimates of classification error decreased, equal to 21.4%, but were still higher than PSF_{breast} OOB estimates of classification error. Variables of importance and their correlations are displayed on Additional file 1: Figure S1b.

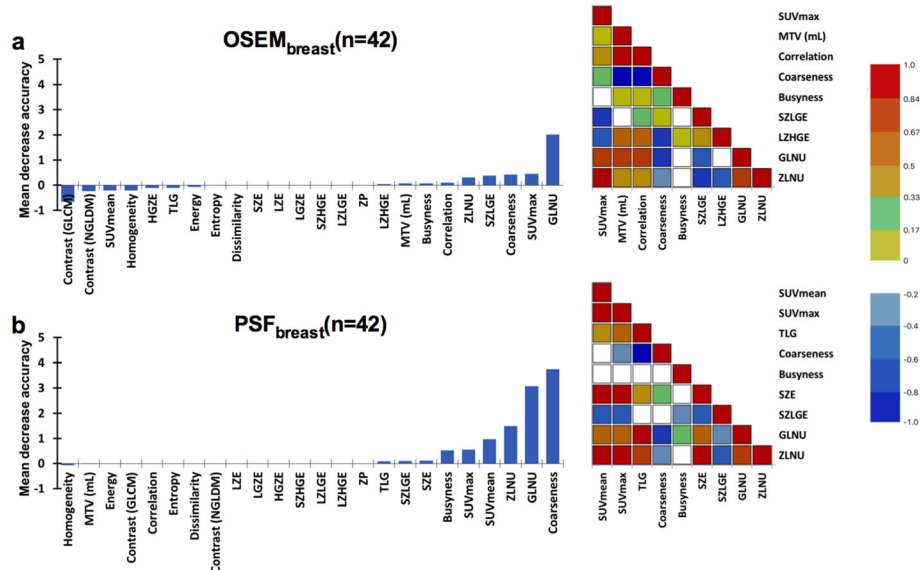


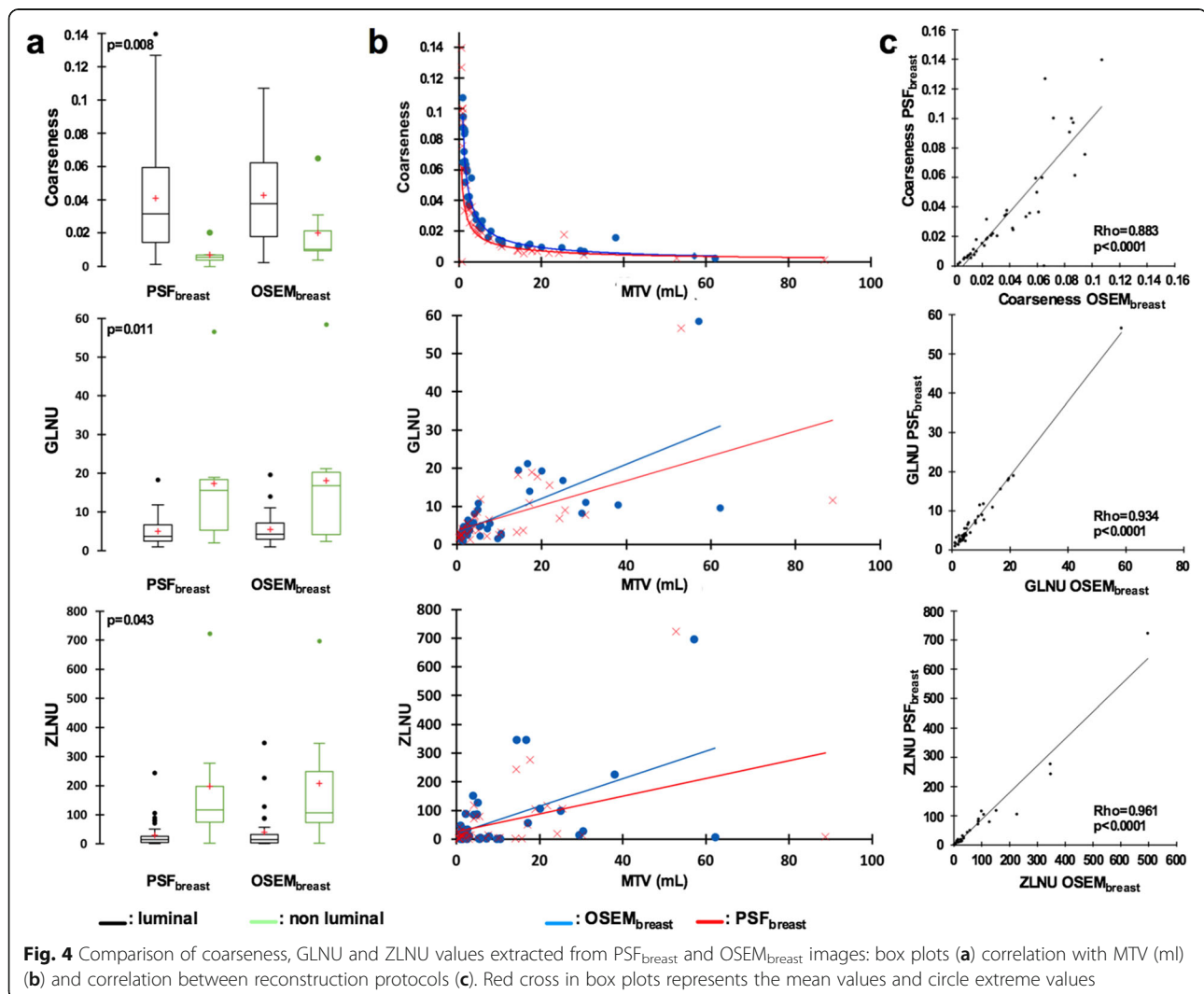
Fig. 3 Left panels display the mean decrease accuracy of textural features values and right panels display Spearman correlation matrixes of all PET metrics found to have positive mean decrease accuracy, whatever the value for OSEM_{breast} (a) and PSF_{breast} (b) reconstructions. For Spearman correlation matrixes, the blue colour corresponds to a correlation close to -1 and the red colour corresponds to a correlation close to 1. The green corresponds to a correlation close to 0

After applying to $\text{PSF}_{\text{breast}}$ images a 4 mm^3 post-reconstruction voxel resampling, the OOB estimates of classification error increased moderately, equal to 26.2%. Variables of importance and their correlations for both protocols are displayed on Additional file 2: Figure S2b. Of note, when applying a 4 mm^3 post-reconstruction voxel resampling on $\text{PSF}_{\text{breast}}$ images, coarseness was then correlated to SUV_{max} values: $\rho = -0.321$, $p = 0.05$.

Comparison of coarseness, GLNU and ZLNU values obtained from $\text{PSF}_{\text{breast}}$ and $\text{OSEM}_{\text{breast}}$ protocols using adapted SUV_{max} bounds for each reconstruction to quantify textural features

Paired comparison of $\text{PSF}_{\text{breast}}$ and $\text{OSEM}_{\text{breast}}$ reconstructions found significant differences between coarseness, GLNU and ZLNU values (Fig. 4a).

Interestingly, the range of coarseness values was wider when using $\text{PSF}_{\text{breast}}$ especially for the smallest lesions, whereas it was quite similar between $\text{PSF}_{\text{breast}}$ and $\text{OSEM}_{\text{breast}}$ for GLNU and ZLNU values (Fig. 4b). However, $\text{PSF}_{\text{breast}}$ and $\text{OSEM}_{\text{breast}}$ coarseness, GLNU and ZLNU values were highly correlated (Fig. 4c). Coarseness displayed the lowest ρ value: 0.883 ($p < 0.0001$) with a dispersion of coarseness values between PET protocols occurring for coarseness values superior to 0.04 corresponding to the smallest lesions ($\text{MTV} < 3 \text{ ml}$). On the contrary, GLNU and ZLNU seem to have the same distribution whatever the protocols and the MTV considered. Moreover, there was no difference between $\text{PSF}_{\text{breast}}$ and $\text{OSEM}_{\text{breast}}$ areas under the ROC for the luminal versus non-luminal status determination with GLNU values and ZLNU values, whereas the area under the ROC



with PSF_{breast} coarseness values was significantly higher than that of $OSEM_{breast}$ coarseness values (Fig. 5). Representative images of one luminal and one non-luminal breast tumours are shown on Fig. 6.

Discussion

As expected, there was a limited number of analysable tumours when using $PSF_{wholeBody}$. Although this reconstruction led to larger MTVs, the number of voxels within MTVs was very low as compared to $OSEM_{breast}$ and PSF_{breast} and thus led to 19 patients (40.4%) being non-exploitable. On the contrary, $OSEM_{breast}$ and PSF_{breast} led to smaller MTVs but a higher number of

voxels and therefore allowed studying nearly all patients: 98% for $OSEM_{breast}$ and 89% for PSF_{breast} .

On matched comparison, PSF_{breast} reconstruction presented better abilities than $PSF_{wholeBody}$ and $OSEM_{breast}$ for the classification of luminal versus non-luminal breast tumours with an accuracy reaching 85.7%. Using the same heterogeneity quantification scale for all three reconstructions, PSF_{breast} still showed higher abilities than others reconstructions. Noticeably, it displayed a high sensitivity but low specificity for the detection of luminal status. Coarseness and ZLNU were the only PET TFs identified as important classification variables by all three reconstruction models. But on $PSF_{wholeBody}$

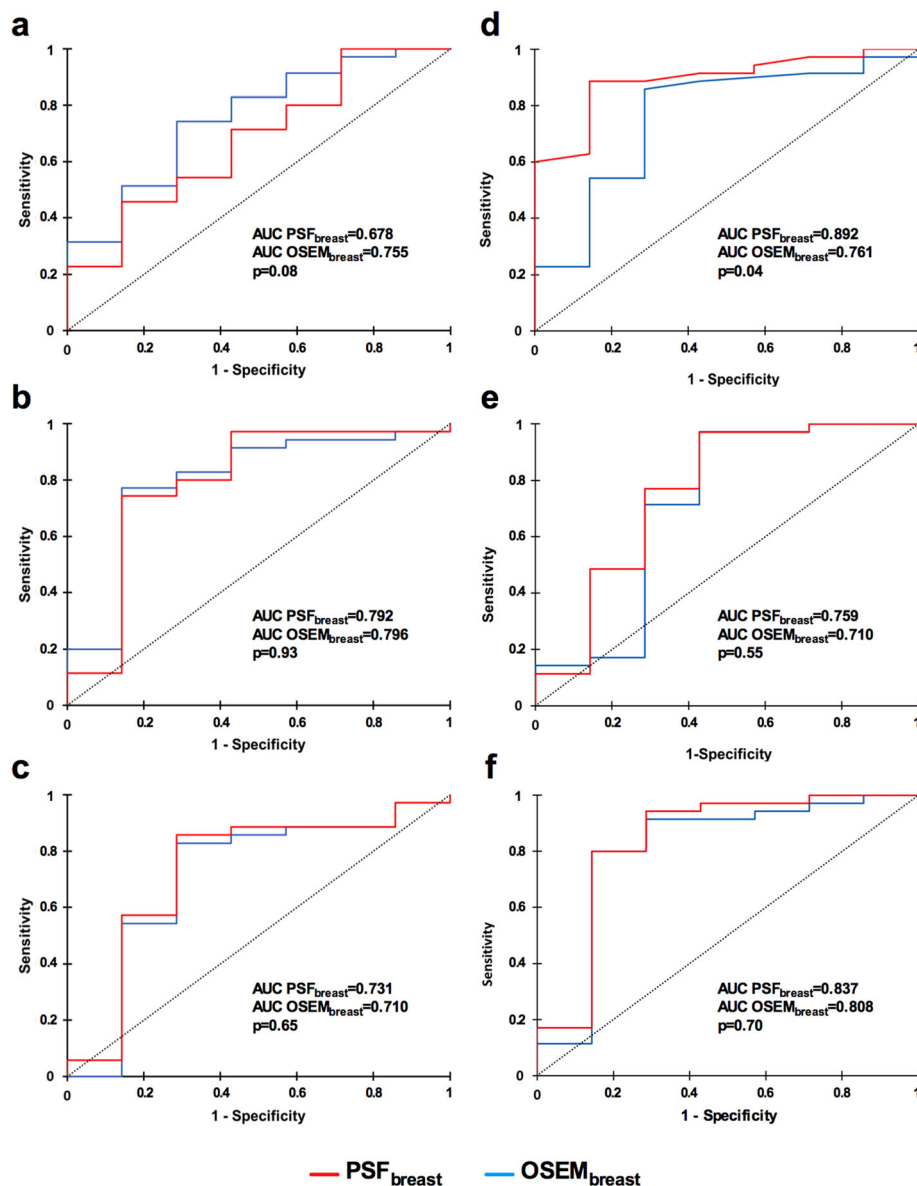
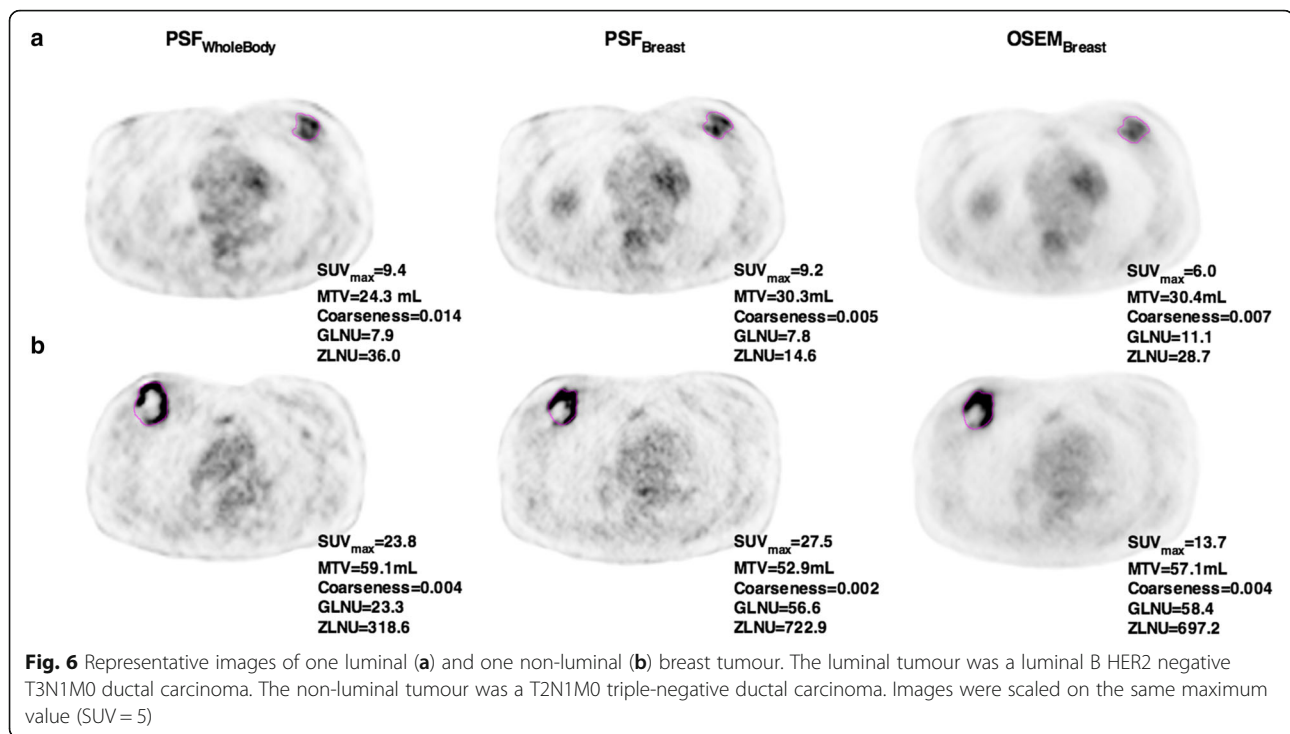


Fig. 5 ROC curves for SUV_{mean} (a), SUV_{max} (b), MTV (c), coarseness (d), GLNU (e) and ZLNU (f) values. The blue line corresponds to $OSEM_{breast}$ and the red one to PSF_{breast} reconstruction



images, coarseness was highly correlated with SUV_{max} , whereas it was not for HR breast protocols. Moreover, correlation between SUV_{max} and coarseness values seems to be linked to voxel size as it disappeared after applying a 2 mm^3 post-reconstruction voxel resampling to $PSF_{wholeBody}$ images and appeared after applying a 4 mm^3 post-reconstruction voxel resampling to PSF_{breast} images. The numerous and strong correlations of TFs with SUV_{max} observed on $PSF_{wholeBody}$ suggested that PET metrics extracted from $PSF_{wholeBody}$ may have less additional information over conventional PET indices. One could consider that the delay between whole body acquisition and the dedicated breast acquisition may have influenced our results. However, this delay was around 20 min and we feel unlikely that it influenced TFs values, as opposed to a previous study in which a second examination was performed 3 h after injection, with a mean time of 127 min between the two phases [35]. Considering image noise, one could have expected that using a small-voxel matrix would have led to higher noise in PSF_{breast} images as compared to $PSF_{wholeBody}$. However, no significant difference in CoV was observed in the present study among all reconstruction protocols, but the small matrix size may have been counterbalanced by a longer acquisition time.

Among HR breast bed position, PSF reconstruction appeared to be more discriminative for luminal versus non-luminal status than OSEM reconstruction. This is in accordance with our previous publication [36] that compared those two types of reconstruction. Regarding

PET metrics extracted from NGLDM matrix and especially coarseness, there was higher values dispersion with PSF_{breast} reconstruction, especially for small lesions and a better area under the ROC for luminal versus non-luminal status determination. Besides, this metric was not correlated to SUV_{max} suggesting that it could provide additional information. Considering TFs extracted from GLZLM, especially ZLNU and GLNU, no difference was found in the dispersion of these TFs values between PSF_{breast} and $OSEM_{breast}$ reconstructions. As coarseness, GLNU and ZLNU were not explored in previous studies, no comparison can be made [17–22].

Concerning heterogeneity quantification process, the main analysis was designed in order to obtain data as close as possible to what could have been done in routine clinical practice, for example in PET units using different reconstruction algorithms. To this end, VOIs and SUV bounds were adapted to each reconstruction independently. To test the influence of quantification scale, a supplemental analysis was made using same SUV bounds for all reconstructions leading to same bin widths and showed no major change as compared to the first analysis. However, as SUV are highly reconstruction-dependent, with for example a mean percentage difference that could reach 66% between OSEM and PSF reconstructions [23], we firmly believe that SUV bounds have to be adapted specifically to the reconstruction of interest. When it comes to VOIs delineation, an appropriate VOI for each reconstruction seems more relevant to answer the question of the influence of reconstruction on FDG radiomics.

Indeed, using the same volume of interest for all reconstructions is never meant to happen in clinical practice. Besides, using same VOIs or independent VOIs between different reconstructions showed almost no influence on a panel of second- and third-order textural features in a previous study [36]. Finally, small-voxel post-reconstruction resampling did not provide better capabilities in terms of histological classification and therefore seems to offer no additional information.

This study had some limitations. First of all, although random forests allowed matched comparison of datasets, it surely did not give definitive results concerning the ability of TFs in discriminating histological characteristics of breast tumours in view of the limited number of patients. The limited number of patients did not allow us to consider all histological tumour subtypes and therefore the discriminative power of TFs was restricted to luminal versus non-luminal tumours. However, the aim of the present study was not to have definitive results concerning PET abilities for histological discrimination. It demonstrated that a combination of PSF modelling and small-voxel reconstruction seems to be the best strategy to obtain additional information over conventional PET metrics and should be used when characterising the intra-tumoral FDG heterogeneity of breast cancers. These results are in line with previous publications using a breast-dedicated PET system, small-voxels and/or new generation reconstruction algorithms with time-of-flight, which found that FDG breast tumour heterogeneity was significantly correlated with immunohistochemical factors and St Gallen's subtypes [18–20], whereas those using OSEM reconstruction with $4 \times 4 \times 4$ mm voxels did not find any association [21, 22].

Conclusions

High-resolution breast PET acquisitions, applying both small-voxel matrix and PSF modelling, appeared to be necessary to improve the characterisation of breast tumours, especially when seeking a link between ^{18}F -fluorodeoxyglucose heterogeneity and histological characteristics in breast cancer.

Additional files

Additional file 1: Figure S1. Impact of quantification scale. Left panels display the mean decrease accuracy of textural features values and right panels display Spearman correlation matrixes of all PET metrics found to have positive mean decrease accuracy, whatever the value for $\text{PSF}_{\text{wholeBody}}$ (a) and $\text{OSEM}_{\text{breast}}$ (b) reconstructions. SUV bounds were set to 0–32 leading to a size of bin of 0.5 for both reconstructions. For Spearman correlation matrixes the blue colour corresponds to a correlation close to -1 and the red colour corresponds to a correlation close to 1. The green corresponds to a correlation close to 0. (TIFF 5689 kb)

Additional file 2: Figure S2. Impact of voxels post-reconstruction resampling. Left panels display the mean decrease accuracy of textural

features values and right panels display Spearman correlation matrixes of all PET metrics found to have positive mean decrease accuracy as well as SUV_{max} and coarseness for $\text{PSF}_{\text{wholeBody}}$ after a 2mm^3 voxels resampling (a) and $\text{PSF}_{\text{breast}}$ after a 4mm^3 voxels resampling (b) reconstructions. For Spearman correlation matrixes the blue colour corresponds to a correlation close to -1 and the red colour corresponds to a correlation close to 1. The green corresponds to a correlation close to 0. (TIFF 6529 kb)

Abbreviations

BMI: Body mass index; CART: Classification and regression trees; CoV: Coefficient of variation; CPP: Comité de Protection des Personnes; DICOM: Digital Imaging and Communications in Medicine; ER: Estrogen receptors; ESMO: European Society for Medical Oncology; FDG: Fluorodeoxyglucose; GLCM: Grey level co-occurrence matrix; GLNU: Grey-level non-uniformity for zone; GLZLM: Grey-level zone length matrix; HER2: Human epidermal growth factor receptor-2; HGZE: Low grey-level zone emphasis; LGZE: Low grey-level zone emphasis; LZE: Long-zone emphasis; LZHGE: Long-zone high grey-level emphasis; LZLGE: Long-zone low grey-level emphasis; MTV: Metabolic tumour volume; NCCN: National comprehensive cancer network; NGLDM: Neighbourhood grey-level different matrix; NPV: Negative predictive value; OOB: Out-of-bag; OSEM: Ordered subset expectation maximization; PET: Positron emission tomography; PPV: Positive predictive value; PR: Progesterone receptors; PSF: Point spread function; Se: Sensitivity; Sp: Specificity; SZE: Short-zone emphasis; SZHGE: Short-zone high grey-level emphasis; SZLGE: Short-zone low grey-level emphasis; TF: Textural feature; VOI: Volume of interest; ZLNU: Zone length non-uniformity; ZP: Zone percentage

Acknowledgements

Not applicable.

Funding

The authors declare that they had no funding.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author upon request.

Authors' contributions

NA designed the study and provided critical readings of the manuscript; TS performed the PET/CT images analysis and helped writing the manuscript; CBF performed the immunohistochemically analyses; JMG collected the clinical data; CLe included the patients; and CLa performed the statistical analysis, wrote the article and built figures and tables. All authors read and approved the final manuscript.

Ethics approval and consent to participate

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.

This study is ancillary to a previous monocentric prospective study conducted by our team and approved by the local Ethics Committee (CPP Nord Ouest III, reference 2009-10). Informed and signed consent was obtained from all patients.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

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

Author details

¹Nuclear Medicine Department, University Hospital, Caen, France. ²INSERM 1199 ANTICIPE, Normandy University, Caen, France. ³Pathology Department,

François Baclesse Cancer Centre, Caen, France. ⁴Biostatistics and Clinical Research Unit, François Baclesse Cancer Centre, Caen, France. ⁵Breast Cancer Unit, François Baclesse Cancer Centre, Caen, France. ⁶Nuclear Medicine Department, François Baclesse Cancer Centre, 3 Avenue du Général Harris, BP 45026 Cedex 5, 14076 Caen, France.

Received: 19 September 2018 Accepted: 10 December 2018

Published online: 29 December 2018

References

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*. 2015;136:E359–86.
2. Turashvili G, Brogi E. Tumor heterogeneity in breast cancer. *Front Med*. 2017;4:227.
3. Groheux D, Espié M, Giacchetti S, Hindié E. Performance of FDG PET/CT in the clinical management of breast cancer. *Radiology*. 2013;266:388–405.
4. Cardoso F, Costa A, Senkus E, Aapro M, André F, Barrios CH, et al. 3rd ESO-ESMO international consensus guidelines for advanced breast cancer (ABC 3). *Breast Edinb Scotl*. 2017;31:244–59.
5. Cochet A, Dygai-Cochet I, Riedinger J-M, Humbert O, Berriolo-Riedinger A, Toubreau M, et al. ¹⁸F-FDG PET/CT provides powerful prognostic stratification in the primary staging of large breast cancer when compared with conventional explorations. *Eur J Nucl Med Mol Imaging*. 2014;41:428–37.
6. Buck A, Schirmer H, Kühn T, Shen C, Kalker T, Kotzerke J, et al. FDG uptake in breast cancer: correlation with biological and clinical prognostic parameters. *Eur J Nucl Med Mol Imaging*. 2002;29:1317–23.
7. Groheux D, Giacchetti S, Moretti J-L, Porcher R, Espié M, Lehmann-Che J, et al. Correlation of high ¹⁸F-FDG uptake to clinical, pathological and biological prognostic factors in breast cancer. *Eur J Nucl Med Mol Imaging*. 2011;38:426–35.
8. Tixier F, Hatt M, Valla C, Fleury V, Lamour C, Ezzouhri S, et al. Visual versus quantitative assessment of intratumor ¹⁸F-FDG PET uptake heterogeneity: prognostic value in non-small cell lung cancer. *J Nucl Med Off Publ Soc Nucl Med*. 2014;55:1235–41.
9. Cook GJR, Yip C, Siddique M, Goh V, Chicklore S, Roy A, et al. Are pretreatment ¹⁸F-FDG PET tumor textural features in non-small cell lung cancer associated with response and survival after chemoradiotherapy? *J Nucl Med Off Publ Soc Nucl Med*. 2013;54:19–26.
10. Desseroit M-C, Visvikis D, Tixier F, Majdoub M, Perdrisot R, Guillemin R, et al. Development of a nomogram combining clinical staging with (18F)-FDG PET/CT image features in non-small-cell lung cancer stage I-III. *Eur J Nucl Med Mol Imaging*. 2016;43:1477–85.
11. Cheng N-M, Fang Y-HD, Chang J-T-C, Huang C-G, Tsan D-L, Ng S-H, et al. Textural features of pretreatment ¹⁸F-FDG PET/CT images: prognostic significance in patients with advanced T-stage oropharyngeal squamous cell carcinoma. *J Nucl Med Off Publ Soc Nucl Med*. 2013;54:1703–9.
12. Tixier F, Le Rest CC, Hatt M, Albarghach N, Pradier O, Metges J-P, 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 Off Publ Soc Nucl Med*. 2011;52:369–78.
13. van Rossum PSN, Fried DV, Zhang L, Hofstetter WL, van Vulpen M, Meijer GJ, et al. The incremental value of subjective and quantitative assessment of ¹⁸F-FDG PET for the prediction of pathologic complete response to preoperative chemoradiotherapy in esophageal cancer. *J Nucl Med Off Publ Soc Nucl Med*. 2016;57:691–700.
14. Bundschuh RA, Dinges J, Neumann L, Seyfried M, Zsótér N, Papp L, et al. Textural parameters of tumor heterogeneity in ¹⁸F-FDG PET/CT for therapy response assessment and prognosis in patients with locally advanced rectal cancer. *J Nucl Med Off Publ Soc Nucl Med*. 2014;55:891–7.
15. Ohri N, Duan F, Snyder BS, Wei B, Machtay M, Alavi A, et al. Pretreatment ¹⁸F-FDG PET textural features in locally advanced non-small cell lung cancer: secondary analysis of ACRIN 6668/RTOG 0235. *J Nucl Med Off Publ Soc Nucl Med*. 2016;57:842–8.
16. Cook GJR, O'Brien ME, Siddique M, Chicklore S, Loi HY, Sharma B, et al. Non-small cell lung cancer treated with Erlotinib: heterogeneity of (18F)-FDG uptake at PET-association with treatment response and prognosis. *Radiology*. 2015;276:883–93.
17. Koolen BB, Vidal-Sicart S, Benlloch Baviera JM, Valdés Olmos RA. Evaluating heterogeneity of primary tumor (18F)-FDG uptake in breast cancer with a dedicated breast PET (MAMMI): a feasibility study based on correlation with PET/CT. *Nucl Med Commun*. 2014;35:446–52.
18. Moscoso A, Ruibal Á, Domínguez-Prado I, Fernández-Ferreiro A, Herranz M, Albaina L, et al. Texture analysis of high-resolution dedicated breast ¹⁸F-FDG PET images correlates with immunohistochemical factors and subtype of breast cancer. *Eur J Nucl Med Mol Imaging*. 2018;45:196–206.
19. Soussan M, Orlhac F, Boubaya M, Zelek L, Zioli M, Eder V, et al. Relationship between tumor heterogeneity measured on FDG-PET/CT and pathological prognostic factors in invasive breast cancer. *PLoS One*. 2014;9:e94017.
20. Antunovic L, Gallivanone F, Sollini M, Sagona A, Invento A, Manfrinato G, et al. [18F]FDG PET/CT features for the molecular characterization of primary breast tumors. *Eur J Nucl Med Mol Imaging*. 2017;44:1945–54.
21. Lemaignier C, Martineau A, Teixeira L, Vercellino L, Espié M, Merlet P, et al. Correlation between tumour characteristics, SUV measurements, metabolic tumour volume, TLG and textural features assessed with ¹⁸F-FDG PET in a large cohort of oestrogen receptor-positive breast cancer patients. *Eur J Nucl Med Mol Imaging*. 2017;44:1145–54.
22. 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 (18F)-FDG PET/CT volumetric and heterogeneity features in stage II/III breast cancer? *Eur J Nucl Med Mol Imaging*. 2015;42:1682–91.
23. Bellevre D, Blanc Fournier C, Switers O, Dugué AE, Levy C, Allouache D, et al. Staging the axilla in breast cancer patients with ¹⁸F-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.
24. Koopman D, van Dalen JA, Arkies H, Oostdijk AHJ, Francken AB, Bart J, et al. Diagnostic implications of a small-voxel reconstruction for loco-regional lymph node characterization in breast cancer patients using FDG-PET/CT. *EJNMMI Res*. 2018;8:3.
25. Koopman D, van Dalen JA, Lagerweij MCM, Arkies H, de Boer J, Oostdijk AHJ, et al. Improving the detection of small lesions using a state-of-the-art time-of-flight PET/CT system and small-voxel reconstructions. *J Nucl Med Technol*. 2015;43:21–7.
26. Boellaard R, O'Doherty MJ, Weber WA, Mottaghy FM, Lonsdale MN, Stroobants SG, et al. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. *Eur J Nucl Med Mol Imaging*. 2010;37:181–200.
27. Lasnon C, Enlhorac B, Popotte H, Aide N. Impact of the EARL harmonization program on automatic delineation of metabolic active tumour volumes (MATVs). *EJNMMI Res*. 2017;7:30.
28. Nioche C, Orlhac F, Boughdad S, Reuzé S, Goya-Outi J, Robert C, et al. LIFEX: a freeware for Radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity. *Cancer Res*. 2018;78:4786–9.
29. Orlhac F, Nioche C, Soussan M, Buvat I. Understanding changes in tumor texture indices in PET: a comparison between visual assessment and index values in simulated and patient data. *J Nucl Med*. 2017;58:387–92.
30. Orlhac F, Soussan M, Chouahnia K, Martinod E, Buvat I. ¹⁸F-FDG PET-derived textural indices reflect tissue-specific uptake pattern in non-small cell lung cancer. *PLoS One*. 10(12):e0145063 [cited 2018 Nov 21]. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4682929/>.
31. Breiman L. Classification and regression trees. New York: Routledge; 1984.
32. Breiman L. Bagging predictors. *Mach Learn*. 1996;24:123–40.
33. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: data mining, inference, and prediction, second edition [internet]. 2nd ed. New York: Springer-Verlag; 2009. [cited 2018 Apr 29]. Available from: www.springer.com/gb/book/9780387848570.
34. Breiman L. Manual on setting up, using, and understanding random forests v4.1. Stat Dep Univ Calif Berkeley. 2002;1. https://www.stat.berkeley.edu/~breiman/Using_random_forests_v4.0.pdf.
35. García-Vicente AM, Molina D, Pérez-Beteta J, Amo-Salas M, Martínez-González A, Bueno G, et al. Textural features and SUV-based variables assessed by dual time point ¹⁸F-FDG PET/CT in locally advanced breast cancer. *Ann Nucl Med*. 2017;31:726–35.
36. Lasnon C, Majdoub M, Lavigne B, Do P, Madelaine J, Visvikis D, et al. ¹⁸F-FDG PET/CT heterogeneity quantification through textural features in the era of harmonisation programs: a focus on lung cancer. *Eur J Nucl Med Mol Imaging*. 2016;43:2324–35.