Annals of Nuclear Medicine (2018) 32:379–388
https://doi.org/10.1007/s12149-018-1253-0

ORIGINAL ARTICLE

Intratumoral heterogeneity in ^{18}F-FDG PET/CT by textural analysis in breast cancer as a predictive and prognostic subrogate

David Molina-García[1] · Ana María García-Vicente[2] · Julián Pérez-Beteta[1] · Mariano Amo-Salas[3] · Alicia Martínez-González[1] · María Jesús Tello-Galán[2] · Ángel Soriano-Castrejón[2] · Víctor M. Pérez-García[1]

Received: 22 February 2018 / Accepted: 1 April 2018 / Published online: 5 June 2018

Abstract

Aim To assess the predictive and prognostic value of textural parameters in locally advanced breast cancer (LABC) obtained by ^{18}F-FDG PET/CT.

Methods Prospective study including 68 patients with LABC, neoadjuvant chemotherapy (NC) indication and a baseline ^{18}F-FDG PET/CT. Breast specimens were grouped into molecular phenotypes and classified as responders or non-responders after completion of NC. Patients underwent standard follow-up to obtain the disease-free survival (DFS) and overall survival (OS). After breast tumor segmentation, three-dimensional (3D) textural measures were computed based on run-length matrices (RLM) and co-occurrence matrices (CM). Relations between textural features with risk categories attending to molecular phenotypes were explored. Kaplan–Meier analysis and univariate and multivariate Cox proportional hazard analysis were used to study the potential of textural variables, molecular phenotypes and histologic response to predict DFS and OS. Receiver operating characteristic (ROC) analysis was used to obtain the best cut-off value, the area under the curve (AUC) and sensitivity and specificity considering OS and DFS.

Results Eighteen patients were classified as responders. Mean ± SD of DFS and OS was 70.87 ± 21.85 and 76.77 ± 18.80 months, respectively. Long run emphasis (LRE) and long run high gray-level emphasis (LRHGE) showed a relation with risk categories. Low gray-level run emphasis (LGRE), LRHGE and run-length non-uniformity (RLNU) showed association with the NC response. Textural variables were significantly associated with OS and DFS in univariate analysis. Regarding the multivariate Cox regression analysis, PET stage with short run high gray-level emphasis (SRHGE) was significantly associated with OS, and PET stage and high gray-level run emphasis (HGRE) with DFS.

Conclusion Textural variables obtained with ^{18}F-FDG PET/CT were predictors of neoadjuvant chemotherapy response and prognosis, being as relevant as PET stage at diagnosis for OS and DFS prediction.

Keywords Breast cancer · ^{18}F-FDG PET/CT · Textural features · Neoadjuvant chemotherapy response · Overall survival · Disease-free survival

Introduction

The quantification of tumor heterogeneity in medical imaging is a current research interest due to its potential relationship with tumor malignancy. High intratumoral heterogeneity has been related to poorer prognosis, which could be secondary to intrinsic aggressive biology or treatment resistance [1, 2]. However, intratumoral heterogeneity is not completely determined by biopsy samples, as they do not reflect the full extent of phenotypic or genetic variability [3].

Information derived from positron emission tomography/computed tomography with ^{18}F-fluorodeoxiglucose (^{18}F-FDG PET/CT) can be a predictor of treatment outcome

✉ Ana María García-Vicente
angarvice@yahoo.es

[1] Mathematical Oncology Laboratory (MôLAB), Universidad de Castilla-La Mancha, Ciudad Real, Spain

[2] Nuclear Medicine Department, University General Hospital, C/ Obispo Rafael Torija s/n, 13005 Ciudad Real, Spain

[3] Department of Mathematics, University of Castilla-La Mancha, Ciudad Real, Spain

[4–7]. In breast cancer (BC), an association between the maximum standardized uptake value (SUVmax) and response has been described. However, results have been controversial [8–12]. On the other hand, limited experience exists about its relation to prognosis [13–16].

Texture analysis refers to a variety of mathematical methods used to quantify the spatial variations in gray levels within an image to derive the so-called 'textural features', which provide a measurement of intralesional heterogeneity. Textural features offer global tumor information and conform approximations of intratumoral heterogeneity, a biological tumor characteristic associated with aggressive tumor behavior, poor response to therapy and poor survival [17]. Summarizing, radiomics support the use of different image characteristics, as textural variables, derived from image-processing techniques and combined with statistical modeling techniques to predict a certain clinical end point (e.g., survival, local relapse or response) [18–22].

Scarcely reported evidence about the relations of textural parameters obtained by ^{18}F-FDG PET/CT with prognostic factors in BC has been described [23]. Moreover, no previous works have studied the relations of textural parameters with neoadjuvant chemotherapy (NC) response.

The aim of the present work was twofold: first, to assess the relations of textural features with risks categories attending to molecular phenotypes, and second, to build a statistical model obtained from 3D texture analysis of PET images identifying the features that can predict NC response and outcome in patients with locally advanced breast cancer (LABC).

Materials and methods

Patients

All reported patients were participants of an ongoing prospective study initiated in September 2009 and approved by the Local Ethics Committee of our Institution and Research Board. Written informed consent was obtained from all patients.

The inclusion criteria were newly diagnosed BC with a size of at least 2 cm in diameter, with clinical indication of NC, lesion uptake higher than background and absence of distant metastases confirmed by previously requested PET/CT for staging purposes. 68 patients satisfied the inclusion criteria.

The histopathological analysis of the primary tumor was performed on specimens obtained by core aspiration biopsy. The determination of tumor type, histopathological grading, estrogen and progesterone receptors (ER/PR), epidermal growth factor receptor (HER2) and proliferation index based on the Ki-67 proliferation ratio were obtained as described in previous works [24–26]. 71.2, 31.8 and 45.5% of patients were HER2, ER and PR negative, respectively, while the remaining patients were positive. TNM was integrated in stages for being more relevant. All patients had ECOG PS of 0. The information of specific adjuvant treatments after NC was not collected.

Additionally, an estimation of the molecular phenotypes and risk categories was performed depending on tumor phenotype: high risk [basal-like or HER2 (+) pure], intermediate risk [luminal B-HER2 (−) or luminal B-HER2(+)] and low risk [luminal A]. Moreover, for statistical analysis, the risk groups were set as categorical variables and divided into two groups: high risk [basal-like and HER2 (+) pure] and low/intermediate risk [luminal B-HER2(-), luminal B-HER2(+) and luminal A] [25].

Patients received standard NC regimen in a combination of anthracyclines, taxanes and anti-HER2 therapy [26].

Sixty patients underwent mastectomy or quadrantectomy and axillary lymph node dissection 4–6 weeks after NC. Breast and lymph nodes specimens were surgically removed, sliced, prepared and analyzed. For this study, only a binary breast histological response was considered to classify lesions as responders (complete or nearly complete response) or non-responders (for the other tumor regression grades) as previously stated [24].

After surgery, adjuvant treatment with/without radiotherapy was administered based on post-NC stage and tumor biology.

Patients underwent a minimum follow-up of 42 months. Disease-free survival (DFS) was defined as the time, in months, from the date of initial staging until tumor recurrence, death or last follow-up examination. Overall survival (OS) was defined as the time, in months, from the date at initial staging until death or last follow-up examination.

FDG PET/CT imaging and tumor segmentation

All PET/CT examinations, previous to NC and surgery, were performed on the same dedicated whole-body PET/CT machine (Discovery DSTE-16 s, GE Medical Systems) following a standardized protocol in three-dimensional (3D) mode [24]. The image voxel size was 5.47 mm × 5.47 mm × 3.27 mm with a slice thickness of 3.27 mm and no gap between slices. Matrix size was 128 × 128.

PET images in DICOM (Digital Imaging and Communication in Medicine) files were imported into the scientific software package Matlab (R2015b, The MathWorks, Inc., Natick, MA, USA) and pre-processed using in-house semi-automatic image segmentation software. The tumor was first manually located in a 3D box and then automatically segmented in three dimensions.

After semi-automatic segmentation, volume of interest was delineated using 40% of SUVmax. SUV-based variables [SUVmax, SUVmean, SUVpeak, metabolic tumor volume (MTV) and total lesion glycolysis (TLG)] were calculated as previously reported [27].

Texture analysis

Many methods have been proposed to quantify tumor heterogeneity from imaging data. For 3D heterogeneity measures, we only considered local and regional spatial textural methods, as we wanted to consider the spatial SUV distribution to assess heterogeneity.

We analyzed the local relations between voxels using the co-occurrence matrix (CM) and the regional relations using the run-length matrix (RLM). Thus, a set of sixteen 3D heterogeneity textural measures was computed automatically using Matlab software [28, 29]. For the computation of the textural features, we considered only the volume of interest segmented (40% of SUVmax). Then, this range was discretized in 32 boxes of equal size to construct the matrices.

Table 1 details the acronyms and description of the assessed textural parameters. Figure 1 shows an example of tumor segmentation and texture image analysis.

Table 1 Textural features and description

Type of measure	Textural features	Description
Co-ocurrence matrix	Entropy (ENT)	Measures the randomness of a gray-level distribution. The ENT is expected to be high if the gray levels are distributed randomly through out the image
	Homogeneity (HOM)	Measures the local homogeneity of a pixel pair. The HOM is expected to be large if the gray levels of each pixel pair are similar
	Contrast (CON)	Measures the local contrast of an image. The CON is expected to be low if the gray levels of each pixel pair are similar
	Dissimilarity (DIS)	Measures the local dissimilarity of an image. The DIS is expected to be low if the gray levels of each pixel pair are similar
	Uniformity (UNI)	Measures the local uniformity of an image. The UNI is expected to be low if the gray levels of each pixel pair are similar
Run-length matrix	Long run emphasis (LRE)	Measures distribution of long runs. The LRE is highly dependent on the occurrence of long runs and is expected large for coarse structural textures
	Short run emphasis (SRE)	Measures the distribution of short runs. The SRE is highly dependent on the occurrence of short runs and is expected large for fine textures
	Low gray-level run emphasis (LGRE)	Measures the distribution of low gray level values. The LGRE is expected large for the image with low gray level values
	High gray-level run emphasis (HGRE)	Measures the distribution of high gray level values. The HGRE is expected large for the image with high gray level values
	Short run low gray-level emphasis (SRLGE)	Measures the joint distribution of short runs and low gray level values. The SRLGE is expected large for the image with many short runs and lower gray level values
	Short run high gray-level emphasis (SRHGE)	Measures the joint distribution of short runs and high gray level values. The SRHGE is expected large for the image with many short runs and high gray level values
	Long run low gray-level emphasis (LRLGE)	Measures the joint distribution of long runs and low gray level values. The LRLGE is expected large for the image with many long runs and low gray level values
	Long run high gray-level emphasis (LRHGE)	Measures the joint distribution of long runs and high gray level values. The LRHGE is expected large for images with many long runs and high gray level values
	Gray-level non-uniformity (GLNU)	Measures the similarity of gray level values through out the image. The GLN is expected small if the gray level values are alike through out the image
	Run-length non-uniformity (RLNU)	Measures the similarity of the length of runs through out the image. The RLN is expected small if the run lengths are alike through out the image
	Run percentage (RPC)	Measures the homogeneity and the distribution of runs of an image in a specific direction. The RPC is the largest when the length of runs is 1 for all gray levels in specific direction

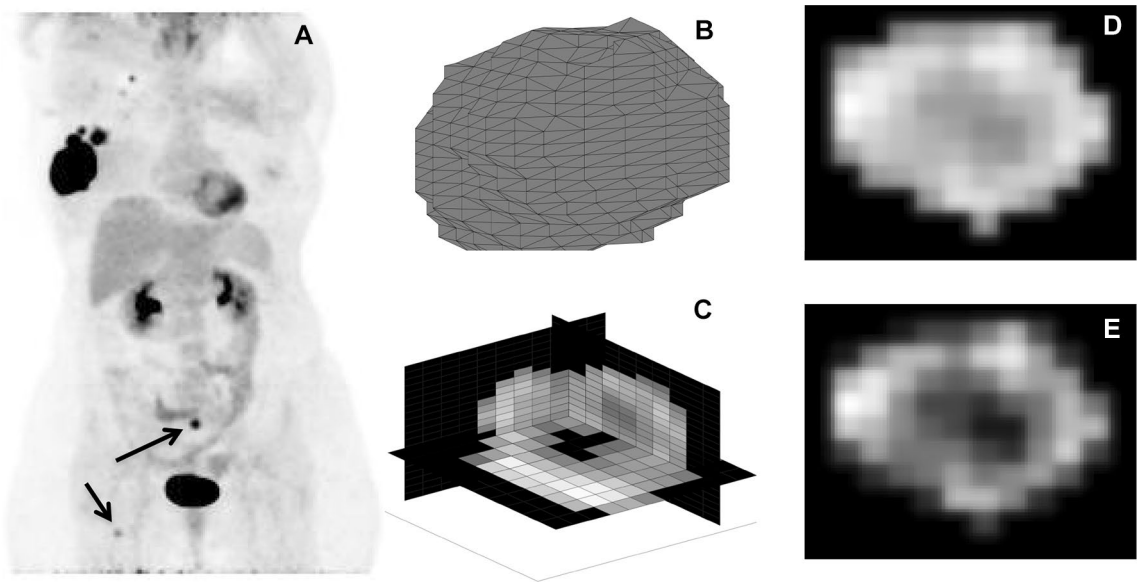

Fig. 1 (**a**) Maximum intensity projection of a patient classified with stage IV due to bone metastases (arrows). Breast tumor segmentation (**b**) and voxel representation in 3D image reconstruction (**c**). Raw gray levels distribution for energy analysis (**d**) and after discretization (**e**)

Statistical methods

The statistical analysis was performed using SPSS software (v. 22.0.00). Categorical variables were described by frequency and percentage, while mean and standard deviation were used to describe quantitative variables. Normality was checked with Kolmogorov–Smirnov test. The level of significance was $p < 0.05$.

Student t-test was used to compare the means of textural variables with respect to NC response groups, using Mann–Whitney test in the nonparametric case.

Relation between phenotype and textural variables was performed using one-way ANOVA, Kruskal–Wallis test in the nonparametric case, and repeated measures ANOVA with a between-subjects factor. Tukey's post hoc test was considered with phenotype variable.

Proportional Cox hazards analysis was used for each textural and SUV-based parameter individually and then in a multivariate analysis where PET stage (distant disease vs regional disease), NC response and molecular phenotype (risk categories) were also included. Hazard ratios were computed to assess the differences in the first significant digit.

Finally, receiver operator characteristic (ROC) curves were used for the computation of the optimal cut-off of the textural features to predict prognosis (DFS or OS) maximizing the sum of sensitivity and specificity. The statistical analysis of survival between groups was performed using Kaplan–Meier survival curves and log-rank test.

Results

68 patients were included. Patient characteristics are detailed in Table 2. Mean ± SD of DFS and OS obtained by Kaplan–Meier analysis were 70.87 ± 21.85 and 76.77 ± 18.80 months, respectively.

Long run emphasis (LRE), LRHGE, LGRE and GLNU showed relations with phenotype risk categories, where higher heterogeneity was related to more aggressive molecular phenotypes (Table 3).

Only LGRE, LRHGE and RLNU showed associations with NC response as shown in Fig. 2. Mean ± SD for LGRE was 0.17 ± 0.04 for responders and 0.20 ± 0.04 for non-responders with $p = 0.048$. LRHGE and RLNU showed marginally significant associations. Mean ± SD for LRHGE was 1178.81 ± 1399.86 for responders and 454.42 ± 374.69 for non-responders with $p = 0.086$. Finally, the mean ± SD values of RLNU for responders was 31.43 ± 21.60 and 20.49 ± 12.30 for non-responders with $p = 0.061$.

Results of the univariate Cox regression analysis used to examine the effects of textural parameters on prognosis are shown in Table 4. CON, DIS, SRE, HGRE, SRHGE and RPC protected against established events (HR < 1), death or recurrence. Then, for example, an increase of 0.10 units of SRE reduced 3.8 and three times the risk of death or recurrence, respectively. On the contrary, HOM, LRE, LRHGE, LRLGE and GLNU increased the risk of death or recurrence (HR > 1). For example, for any 0.10 units of increase of HOM, the risk of death or recurrence increased 15.6 or

Table 2 Patient's characteristics

Characteristics	*n*	%
Histology		
IDC	65	95.5
ILC	3	4.5
Risk phenotype		
Low	5	7.3
Intermediate	41	60.3
High	22	32.4
NC breast response		
Yes	18	26.5
No	43	63.2
n.a.	7	10.3
Metabolic stage		
II or III	56	82.4
IV	12	17.6
Recurrence		
Yes	10	14.7
No	56	82.3
No disease free	2	3.0
Death		
Yes	9	13.2
No	59	86.8

NC neoadjuvant chemotherapy, *n* number of patients

Table 3 Textural variables depending on the risk phenotype categories (*)

Textural variables	Risk categories	mean ± SD values	*p* values
LRE	Low–Intermediate	9.27 ± 7.23	**0.027**
	High	25.66 ± 41.54	
LRHGE	Low–Intermediate	432.42 ± 341.27	**0.021**
	High	933.50 ± 1187.02	
LGRE	Low–Intermediate	0.19 ± 0.04	0.082
	High	0.18 ± 0.03	
GLNU	Low–Intermediate	5.06 ± 3.57	0.065
	High	7.57 ± 5.99	

Significant values are represented in bold

(*) Only textural variables that showed significant or close to the significant association with risk categories (as categorical variable) are included

SD standard deviation

9.4 times, respectively. With respect to LRHGE, the risk of experiencing an event increased by 7% for every increase of 100 units of the variable.

Kaplan–Meier analysis showed that OS and DFS were lower in patients with high HOM, SRE or RPC. For HOM, values lower than the most significant threshold had higher survival. On the contrary, higher values of SRE and RPC were associated with lower survival. (Figs. 3, 4 and 5).

The best cut-off value, area under curve (AUC) and sensitivity and specificity obtained in ROC analysis for the OS prediction were, respectively, 2.68, 0.808, 75 and 88% for HOM ($p = 0.005$); 5.90, 0.756, 88 and 67% for SRE ($p = 0.019$), and 4.69, 0.804, 88 and 67% for RPC ($p = 0.006$). For DFS, results were: 2.43, 0.770, 67 and 74% for HOM ($p = 0.010$); 5.90, 0.703, 78, 67% for SRE ($p = 0.052$) and 4.84, 0.725, 78, 62% for RPC ($p = 0.031$).

In the multivariable Cox regression analysis including SUV-based and textural variables, PET stage (distant disease vs regional disease), NC response and risk categories (molecular phenotype), only two variables showed prognostic value. For OS, PET stage and SRHGE showed significant associations. Regarding PET stage, a patient with distant metastases had 12.5 times greater risk of death compared to a patient with regional disease (HR 12.50; 95% IC of 2.58 and 60.41; $p = 0.002$). On the contrary, SRHGE was a protective variable (HR 0.76; 95% IC of 0.62 and 0.92; $p = 0.006$). For DFS, only PET stage (HR 4.72; 95% IC of 1.21 and 18.37, $p = 0.025$) and HGRE (HR 0.89; 95% IC of 0.80 and 0.99, $p = 0.035$) showed significant associations.

Discussion

Several textural features, in functional imaging of cancer using PET, have shown ability to differentiate tumor types, predict treatment response, or be associated with survival [7, 30–32].

In BC, more aggressive tumors are associated with a higher metabolism, tumor burden and also a better histological response to NC [25, 33]. Yoon et al. [34] evaluated the intratumoral metabolic heterogeneity by constructing cumulative SUV histograms. They concluded that lower AUC of those histograms was correlated with the invasive component of breast ductal carcinoma in situ. Soussan et al. [23] found associations between textural features and negative ER, negative PR and triple-negative breast cancer, concluding that tumor heterogeneity assessed on ^{18}F-FDG PET/CT might be used to determine breast cancer aggressiveness.

In this work, we analyzed the relation between textural parameters and histopathological prognostic factors combined into phenotype risk categories. Textural features, as LRHGE, LRE and GLNU, showed relations with risk phenotypes, pointing out a relationship between local tumor aggressiveness and tumor complexity. Variables based on co-occurrence matrix were not good predictors, neither for risk phenotype or for NC response. Both LGRE and LRHGE showed significant differences between risk phenotypes and NC response.

Our analysis presents similarities with the one developed by Soussan et al. [23] who first used 3D matrices to construct a set of textural measures, studied their

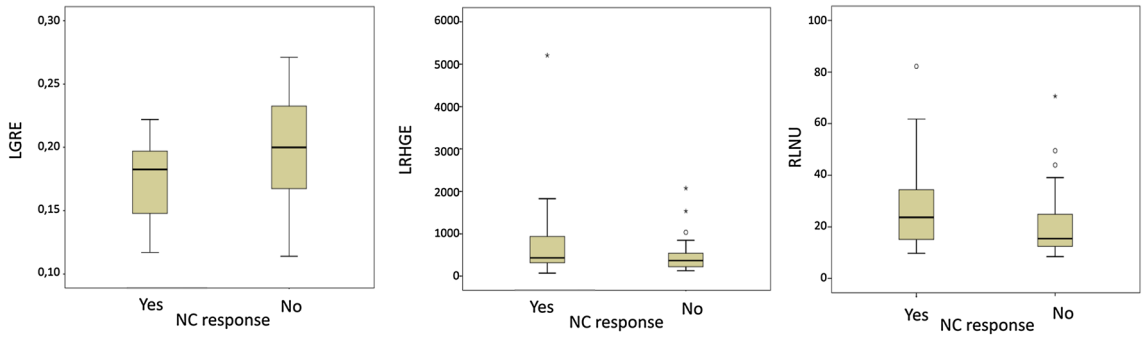

Fig. 2 Texture variables associated with neoadjuvant chemotherapy response. *p* values of 0.048, 0.086 and 0.061 were obtained for LGRE, LRHGE and RLNU, respectively

Table 4 Cox regression univariate analysis for textural features for OS and DFS

SUV-based variables	OS		DFS	
	p value	HR (95% CI)	*p* value	HR (95% CI)
SUV max	0.983	0.99 (0.88–1.13)	0.532	1.04 (0.93–1.15)
SUV mean	0.855	0.98 (0.80–1.20)	0.706	1.04 (0.87–1.23)
SUV peak	0.996	1.00 (0.86–1.16)	0.576	1.04 (0.91–1.19)
MTV	***0.011***	1.03 (1.01–1.05)	***0.023***	1.02 (1.01–1.04)
TLG	***0.049***	1.002 (1.00–1.004)	***0.058***	1.002 (1.00–1.004)
Textural variables (*)	*p* value	HR (95% CI)	*p* value	HR (95% CI)
ENT	0.189	0.14 (0.01–2.67)	0.383	0.28 (0.02–4.93)
HOM	***<0.001***	15.63 (3.37–72.52)	***0.001***	9.42 (2.42–36.63)
CON	***0.028***	0.89 (0.80–0.99)	0.062	0.92 (0.85–1.00)
DIS	***0.009***	0.21 (0.06–0.67)	***0.023***	0.32 (0.12–0.85)
UNI	0.056	1.2 (0.99–1.44)	0.117	1.16 (0.96–1.39)
SRE	***0.019***	0.26 (0.08–0.80)	***0.025***	0.32 (0.12–0.87)
LRE	***<0.001***	1.02 (1.01–1.04)	***0.002***	1.02 (1.01–1.03)
LGRE	0.108	4.52 (0.72–28.46)	0.180	3.13 (0.59–16.52)
HGRE	***0.007***	0.84 (0.75–0.96)	***0.012***	0.88 (0.79–0.97)
SRLGE	0.693	1.42 (0.25–7.93)	0.878	1.13 (0.23–5.69)
SRHGE	**0.005**	0.81 (0.70–0.94)	***0.031***	0.89 (0.80–0.99)
LRLGE	**0.003**	1.10 (1.04–1.18)	***0.007***	1.09 (1.02–1.16)
LRHGE	**0.005**	1.07 (1.02–1.12)	**0.021**	1.06 (1.01–1.11)
GLNU	**0.009**	1.14 (1.03–1.25)	**0.020**	1.12 (1.02–1.23)
RLNU	0.052	1.03 (1.00–1.07)	0.086	1.03 (0.99–1.06)
RPC	***0.004***	0.38 (0.20–0.74)	***0.016***	0.50 (0.28–0.88)

Significant values are represented in bold and bolditalics

OS overall survival, *DFS* disease-free survival, *HR* hazard ratio, *CI* confidence interval

(*) Textural variables are explained in Table 1

association with biological features and found that the combination of HGRE and SUVmax identified triple-negative tumors with a sensitivity of 77% and specificity of 71%. On the contrary, Groheux et al. [35] found that none of the considered PET texture metrics could improve differentiation between the three main molecular subtypes of breast tumors beyond the standard clinical factors and SUV metrics.

Tixier et al. [7], in patients with esophageal carcinoma, found that responders to treatment showed greater local and regional heterogeneity at baseline, with better response stratification for the measures of regional tumor heterogeneity.

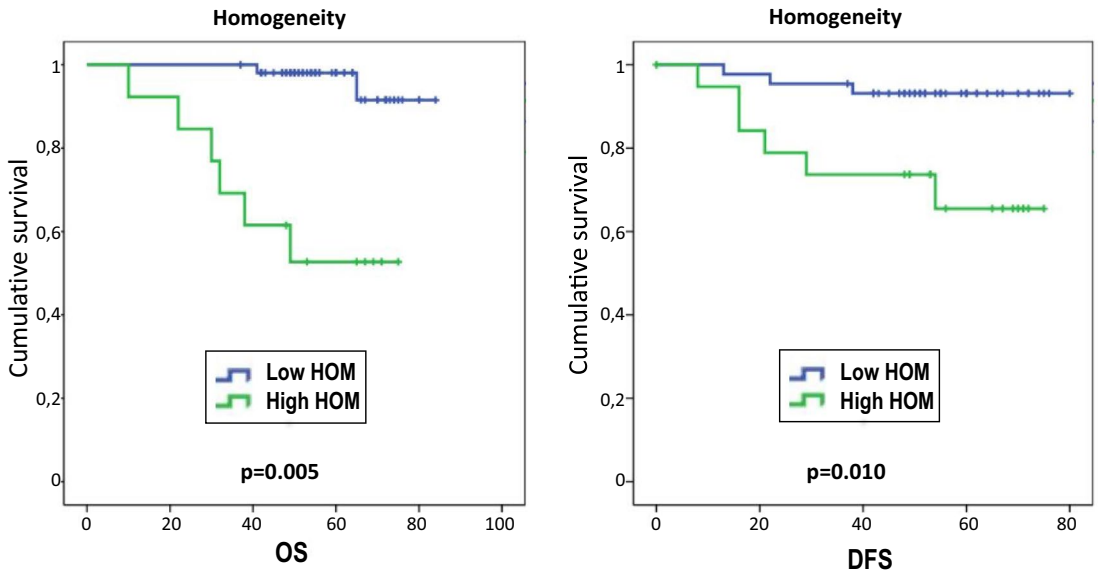

Fig. 3 Kaplan–Meier plots for OS and DFS for patient subgroups with different values of the textural variable "homogeneity" (HOM). Log-rank *p* values are provided

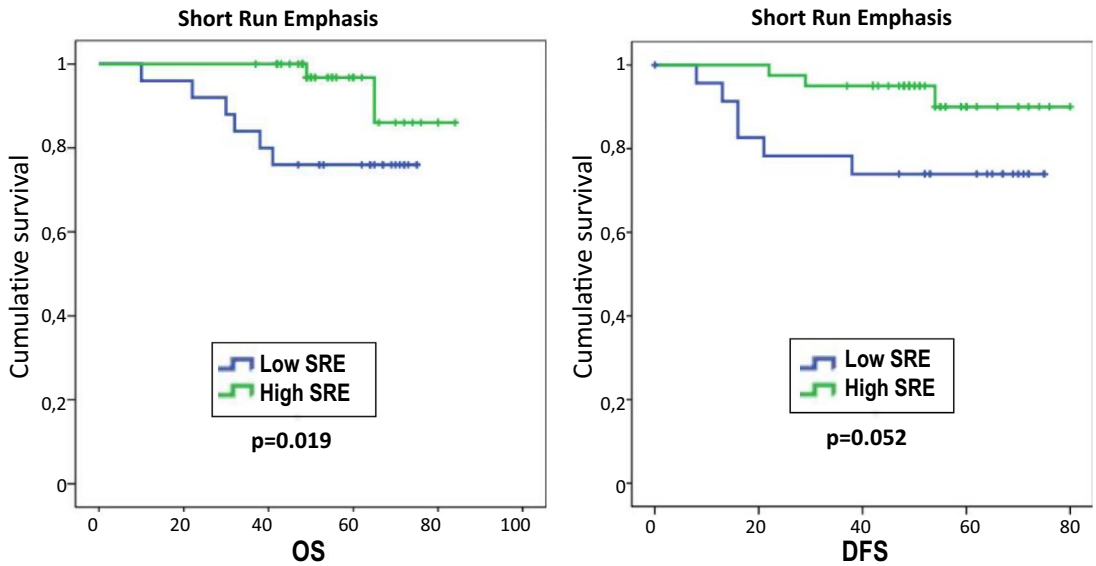

Fig. 4 Kaplan–Meier plots for OS and DFS for patient subgroups with different values of the textural variable "short run emphasis" (SRE). Log-rank *p* values are provided

Our results showed similarities, since more heterogeneous tumor values had a better NC response. The fact that not all the textural parameters were associated with response has been previously noted [36]. In the present work, only LGRE, LRHGE and RLNU showed significant or marginally significant associations. Thus, lesions with high gray level values (LGRE, LRHGE) and with a more heterogeneous distribution of the radiotracer (RLNU) had a better NC response compared to the other groups.

Previously reported experience about the association between ^{18}F-FDG tumor uptake and prognosis has outlined controversial results, probably due to the differences of tumor biology and methodology among the works [37–39]. Thus, SUVmax does not seem to be a strong variable to predict disease evolution. Regarding volume-based variables obtained in baseline PET, there is less evidence with respect to their prognostic value [13–16]. The addition of MTV to other variables, as tumor biology, has potential benefits for

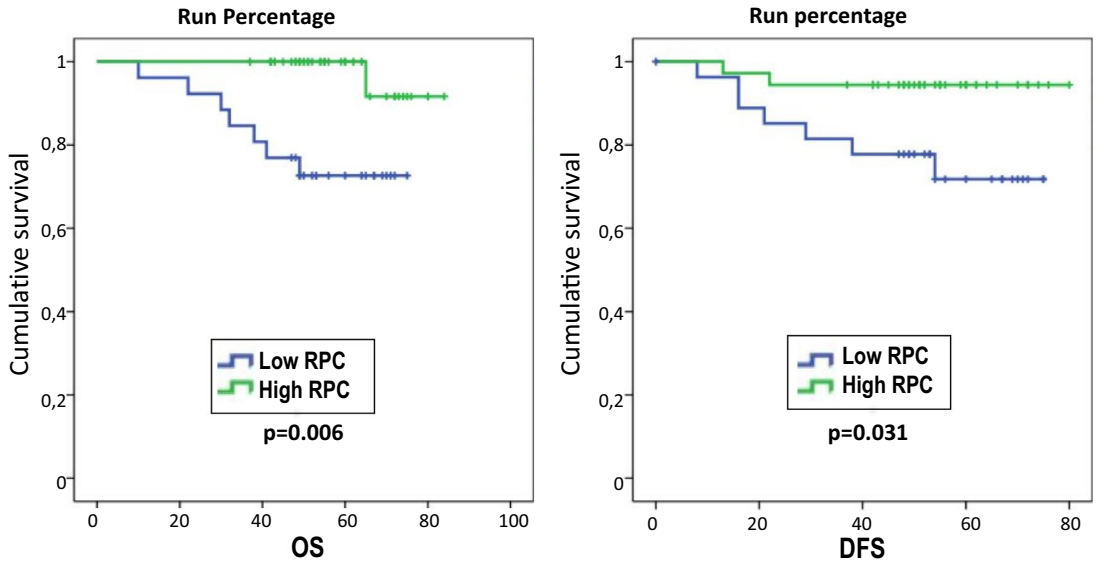

Fig. 5 Kaplan–Meier plots for OS and DFS for patient subgroups with different values of the textural variable "run percentage" (RPC). Log-rank *p* values are provided

identifying a subgroup of patients at higher risk for recurrence, although the results are controversial [13, 40].

The mainstay of radiomics is to build clinical models to predict patient outcome, thereby facilitating better patient management. Intratumoral heterogeneity assessed by PET has been described as an independent prognostic factor in several tumors [41–43]. In patients with non-small cell lung cancer, both heterogeneity and tumor size were predictive for disease-specific survival, but only texture determined by CM Entropy was determined as an independent factor in multivariate analysis. On the other hand, OS was not significantly correlated, most likely due to the high comorbidity in the cohort [41]. In pancreatic ductal adenocarcinoma, Hyun et al. [42] found that intratumoral heterogeneity of ^{18}F-FDG uptake was an independent survival prognostic factor. However other authors have found no significant associations with prognosis in multivariate analysis [32, 38]. Cook et al. [32] found that progression-free survival was longer in patients with high levels of contrast and busyness, although they did not find an association with OS. Our results are in accordance with these and may be explained by the fact that some textural variables were not representative of tumor heterogeneity. Moreover, the connection between biology and treatment response could influence the survival results, with more biologically aggressive tumors showing a better NC response and thus prognosis. However, we did not find a significant association between prognosis and treatment response, probably due the low number of responders.

In BC, limited reported evidence exists regarding texture and prognosis. Previous authors have described that textural features performed better than SUV parameters, MTV and TLG in the determination of prognosis [16]. Son et al. [43] found an association between a heterogeneity volume-based parameter and OS.

Textural features give information about gray level distribution and the image-visible heterogeneity. On the other hand, not all the texture variables underlay information about tumor heterogeneity in the same way or have the same potential in the heterogeneity description. Some variables offer a distinction between fine textures and coarse ones. Co-occurrence matrix variables give less robust information of texture compared to RLM variables. Thus, the obtained results with CM variables should be cautiously considered.

The results of our prognostic model revealed that for a constant value of SRHGE, PET stage (distant metastases) was a strong predictor for OS. On the other hand, the SRHGE was the best texture variable in the OS prediction. The information of the high gray level voxel distribution supporting more robust information of tumor heterogeneity was comparable to other texture variables. However, when SRHGE was not included in the model, other texture variables appeared to be relevant. This fact can explain the overlapping information offered by some texture variables. For DFS, similar behavior was observed between PET stage and HGRE. Thus, texture and PET stage overcame SUV-based variables and tumor biology.

About the limitations of the study, the reduced and biologically heterogeneous sample could affect the obtained results, although we considered it as representative of the total population of patients included in our prospective study. Moreover, the results obtained were focused on locally advanced tumors and thus, could not be applicable

to smaller tumors. Regarding tumor volume, although a selection of tumors with a size of at least 2 cm of diameter was performed, textural variables could not reproduce real tumor heterogeneity, based on the limited number of analyzed voxels.

Regarding the strengths, patients were normalized in a prospective study, tumors were segmented and textural parameters were computed in 3D, using a refined methodology.

The results obtained in this work offer evidence of the associations between textural parameters, treatment response and prognosis in breast cancer.

Conclusions

Texture variables obtained with ^{18}F-FDG PET/CT, were predictors of neoadjuvant chemotherapy response and prognosis, being almost as relevant as PET stage at diagnosis for OS and DFS prediction.

Normalized and well-computed heterogeneity parameters obtained by PET combined with the metabolic stage can help in the assessment of patient prognosis.

Acknowledgements This work has been supported by Ministerio de Economía y Competitividad/FEDER, Spain [Grant number MTM2015-71200-R] and Consejería de Educación Cultura y Deporte from Junta de Comunidades de Castilla-La Mancha, Spain [Grant number PEII-2014-031-P].

Funding This study was funded by Ministerio de Economía y Competitividad/FEDER, Spain [Grant number MTM2015-71200-R] and Consejería de Educación Cultura y Deporte from Junta de Comunidades de Castilla-La Mancha, Spain [Grant number PEII-2014-031-P].

Compliance with ethical standards

Conflict of interest The authors declare that they do not have any conflicts of interest related to this work.

Human participants and animal rights statement 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 obtained from all individual participants included in the study.

References

1. Fischer R, Pusztai L, Swanton C. Cancer heterogeneity: implications for targeted therapeutics. Br J Cancer. 2013;108:479–85.
2. Bonavia R, Inda M, Cavenee W, Furnari F. Heterogeneity maintenance in glioblastoma: asocial network. Cancer Res. 2011:12: 4055–60.
3. Aparicio S, Caldas C. The implications of clonal genome evolution for cancer medicine. N Engl J Med. 2013;9:842–51.
4. Eary JF, O'Sullivan F, O'Sullivan J, Conrad EU. Spatial heterogeneity in sarcoma 18F-FDG uptake as a predictor of patient outcome. J Nucl Med. 2008;49:1973–79.
5. El Naqa I, Grigsby PW, Aptea 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.
6. Yu H, Caldwell C, Mah K, Poon I, Balogh J, MacKenzie R, et al. Automated radiation targeting in head-and-neck cancer using region-based texture analysis of PET and CT images. Int J Radiat Oncol Biol Phys. 2009;75:618–25.
7. Tixier F, Le Rest CC, Hatt M, Albarghach N, Pradier O, Metges JP, et al. Intratumor heterogeneity characterized by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. J Nucl Med. 2011;52:369–78.
8. Schwarz-Dose J, Untch M, Tiling R, Sassen S, Mahner S, Kahlert S, et al. Monitoring primary systemic therapy of large and locally advanced breast cancer by using sequential positron emission tomography imaging with [18F] fluorodeoxyglucose. J Clin Oncol. 2009;27:535–41.
9. Buchbender C, Kuemmel S, Hoffmann O, Stahl AR, Kimmig R, Otterbach F, et al. FDG-PET/CT for the early prediction of histopathological complete response to neoadjuvant chemotherapy in breast cancer patients: initial results. Acta Radiol. 2012;53:628–36.
10. Ueda S, Saeki T, Shigekawa T, Omata J, Moriya T, Yamamoto J, et al. 18F-Fluorodeoxyglucose positron emission tomography optimizes neoadjuvant chemotherapy for primary breast cancer to achieve pathological complete response. Int J Clin Oncol. 2012;17:276–82.
11. Berriolo-Riedinger A, Touzery C, Riedinger JM, Toubeau M, Coudert B, Arnould L, et al. [18F]FDG-PET predicts complete pathological response of breast cancer to neoadjuvant chemotherapy. Eur J Nucl Med Mol Imaging. 2007;34:1915–24.
12. Kolesnikov-Gauthier H, Vanlemmens L, Baranzelli MC, Vennin P, Servent V, Fournier C, et al. Predictive value of neoadjuvant chemotherapy failure in breast cancer using FDG-PET after the first course. Breast Cancer Res Treat. 2012;131:517–25.
13. Kim J, Yoo SW, Kang SR, Cho SG, Oh JR, Chong A, et al. Prognostic significance of metabolic tumor volume measured by (18) F-FDG PET/CT in operable primary breast cancer. Nucl Med Mol Imaging. 2012;46:278–85.
14. Nakajima N, Kataoka M, Sugawara Y, Ochi T, Kiyoto S, Ohsumi S, et al. Volume-based parameters of 18F-fluorodeoxyglucose positron emission tomography/computed tomography improve disease recurrence prediction in postmastectomy breast cancer patients with 1 to 3 positive axillary lymph nodes. Int J Radiat Oncol Biol Phys. 2013;87:738–46.
15. Yue Y, Cui X, Bose S, Audeh W, Zhang X, Fraass B. Stratifying triple-negative breast cancer prognosis using 18F-FDG-PET/CT imaging. Breast Cancer Res Treat. 2015;153:607–16.
16. Chen S, Ibrahim NK, Yan Y, Wong ST, Wang H, Wong FC. Risk stratification in patients with advanced-stage breast cancer by pretreatment [(18)F]FDG PET/CT. Cancer. 2015;121:3965–74.
17. Shin Y, Han S, Chung E, Chung S. Intratumoral phenotypic heterogeneity as an encourager of cancer invasion. Integr Biol (Camb). 2014;6:654–61.
18. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer. 2012;48:441–6.
19. Larue RT, Defraene G, De Ruysscher D, Lambin P, van Elmpt W. Quantitative radiomics studies for tissue characterization: a

review of technology and methodological procedures. Br J Radiol. 2017;90(1070):20160665. https://doi.org/10.1259/bjr.20160665.
20. Lian C, Ruan S, Denoeux T, Jardin F, Vera P. Selecting radiomic features from FDG-PET images for cancer treatment outcome prediction. Med Image Anal. 2016;32:257–68.
21. Li LM, Castellano C, Bonilha L, Cendes F. Texture analysis of medical images. Clin Radiol. 2004;59:1061–69.
22. Chicklore S, Goh V, Siddique M, Roy A, Marsden PK, Cook GJ. Quantifying tumour heterogeneity in 18F-FDG PET/CT imaging by texture analysis. Eur J Nucl Med Mol Imaging. 2013;40:133 – 40.
23. Soussan M, Orlhac F, Boubayaand M, Zelek L, Ziol M, Eder V, et al. Relationship between tumor heterogeneity measured on FDG-PET/CT and pathological prognosis factors in invasive breast cancer. PLoS One. 2014;9:e94017.
24. García Vicente AM, Cruz Mora MA, León Martín AA, Muñoz Sanchez MM, Relea Calatayud F, Van Gómez López O, et al. Glycolitic activity with 18F-FDG PET/CT predicts final neoadjuvant chemotherapy response in breast cancer. Tumour Biol. 2014;35:11613–20.
25. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thürlimann B, Senn HJ, et al. Strategies for subtypes-dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. Ann Oncol. 2011;22:1736–47.
26. García Vicente AM, Soriano Castrejón A, León Martín A, Relea Calatayud F, Muñoz Sánchez MM, Cruz Mora MA, et al. Early and delayed prediction of axillary lymph node neoadjuvant response by ^{18}F-FDG PET/CT in patients with locally advanced breast cancer. Eur J Nucl Med Mol Imaging. 2014;41:1309–18.
27. García Vicente AM, Pérez-Beteta J, Pérez-García VM, Molina-García D, Jiménez Londoño GA, Soriano Castrejón A, Martínez-González A. Metabolic tumor burden assessed by dual time point 18F-FDG PET/CT in locally advanced breast cancer: relation with tumor biology. Mol Imaging Biol. 2017;19:636–44.
28. Molina D, Pérez-Beteta J, Martínez-González A, Martino J, Velásquez C, Arana E, Pérez-García VM. Influence of gray level and space discretization on brain tumor heterogeneity measures obtained from magnetic resonance images. Comput Biol Med. 2016;78:49–57.
29. Molina D, Pérez-Beteta J, Martínez-González A, Sepúlveda JM, Peralta S, Gil-Gil MJ, et al. Geometrical Measures Obtained from Pretreatment Postcontrast T1 Weighted MRIs Predict Survival Benefits from Bevacizumab in Glioblastoma Patients. PLoS One. 2016;11:e0161484.
30. Cheng NM, Fang Y-H, Lee L, Chang J, Tsan D-L, Ng S-H, et al. Zone-size nonuniformity of 18F-FDG PET regional textural features predicts survival in patients with oropharyngeal cancer. Eur J Nucl Med Mol Imaging. 2015;42:419–28.
31. Pyka T, Gempt J, Hiob D, Ringel F, Schlegel J, Bette S, et al. Textural analysis of pre-therapeutic [18F]-FET-PET and its correlation with tumor grade and patient survival in high-grade gliomas. Eur J Nucl Med Mol Imaging. 2016;43:133–41.
32. Cook G, 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. 2013;54:19–26.
33. Cortazar P, Zhang L, Untch M, Mehta K, Costantino JP, Wolmark N, et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. Lancet 2014;384:164–72.
34. Yoon H, Kim Y, Kim BS. Intratumoral metabolic heterogeneity predicts invasive components in breast ductal carcinoma in situ. Eur Radiol. 2015;12:3648–58.
35. 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.
36. Nakajo M, Jinguji M, Nakabeppu Y, Nakajo M, Higashi R, Fukukura Y, et al. Texture analysis of 18F-FDG PET/CT to predict tumour response and prognosis of patients with esophageal cancer treated by chemoradiotherapy. Eur J Nucl Med Mol Imaging. 2017;44:206–14.
37. García Vicente AM, Soriano Castrejón A, López-Fidalgo JF, Amo-Salas M, Muñoz Sanchez MM, Álvarez Cabellos R, et al. Basal ^{18}F-FDG PET/CT as a prognostic biomarker in patients with locally advanced breast cancer. Eur J Nucl Med Mol Imaging. 2015;42:1804–13.
38. O JH, Choi WH, Han EJ, Choi EK, Chae BJ, Park YG, et al. The prognostic value of 18F-FDG PET/CT for early recurrence in operable breast cancer: comparison with TNM Stage. Nucl Med Mol Imaging. 2013;47:263–67.
39. Inoue T, Yutani K, Taguchi T, Tamaki Y, Shiba E, Noguchi S. Preoperative evaluation of prognosis in breast cancer patients by [(18)F]2-Deoxy-2-fluoro-D-glucose-positron emission tomography. J Cancer Res Clin Oncol. 2004;130:273–8.
40. 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.
41. Pyka T, Bundschuh RA, Andratschke N, Mayer B, Specht HM, Papp L, et al. Textural features in pre-treatment [F18]-FDG-PET/CT are correlated with risk of local recurrence and disease-specific survival in early stage NSCLC patients receiving primary stereotactic radiation therapy. Radiat Oncol. 2015;10:100.
42. Hyun SH, Kim HS, Choi SH, Choi DW, Lee JK, Lee KH, et al. Intratumoral heterogeneity of 18F-FDG uptake predicts survival in patients with pancreatic ductal adenocarcinoma. Eur J Nucl Med Mol Imaging. 2016;43:1461–68.
43. Son SH, Kim DH, Hong CM, Kim CY, Jeong SY, Lee SW, et al. Prognostic implication of intratumoral metabolic heterogeneity in invasive ductal carcinoma of the breast. BMC Cancer. 2014;14:585–96.