Journal of Nuclear Medicine, published on June 3, 2016 as doi:10.2967/jnumed.116.173708

Multi-scale texture analysis: from 18F-FDG PET images to pathological slides.

Authors: Fanny Orlhac[1] (PhD), Benoit Thézé[1] (MSc), Michaël Soussan[1,2] (MD,PhD), Raphaël Boisgard[1] (PhD), Irène Buvat[1] (PhD)

1: Imagerie Moléculaire In Vivo, Inserm, CEA, Univ. Paris Sud, CNRS, Université Paris Saclay, CEA - Service Hospitalier Frédéric Joliot, Orsay, France

2: Department of Nuclear Medicine, AP-HP, Avicenne Hospital, Bobigny, France.

Corresponding author: Fanny Orlhac, PhD

4 place du Général Leclerc

91400 Orsay

France

Tel: 33 1 69 86 78 21

Fax: 33 1 69 86 77 86

Email: orlhacf@gmail.com

Word count: 4,998 words

ABSTRACT

Purpose: Characterizing tumor heterogeneity using texture indices (TI) calculated from Positron Emission Tomography (PET) images has shown promise in predicting treatment response or patient survival in some types of cancer. Yet, the relationship between PET-derived TI, precise tracer distribution and biological heterogeneity needs to be clarified. We investigated the relationship between PET-derived TI and observations made on autoradiography and histological slides.

Methods: Three mice bearing orthopically implanted mammary tumors derived from transgenic MMTV-PyMT mouse were scanned with the Inveon PET/Computerized Tomograph after injection of 18F-FDG. Tumors were then sliced, autoradiography was performed and tumor slices were stained using haematoxylin and eosin. Six TI computed from PET, autoradiography and histological images were compared to assess the ability of texture analysis to capture heterogeneity at different scales.

Results: TI were significantly correlated (Spearman coefficient R between 0.57 and 0.85) between PET and autoradiography images but the TI values differed in magnitude. The TI correlation was low between histological and PET or autoradiography images (R between 0.06 and 0.54). All TI were little or moderately influenced by the difference of voxel size and spatial resolution in autoradiography images. TI measured in autoradiography images were significantly different ($p<0.05$) between histological regions with a high density of cells and regions with a low cell density, or between regions presenting different spatial arrangements of cells.

Conclusion: Heterogeneity computed *in vivo* from PET images accurately reflects the heterogeneity of tracer uptake directly computed *ex vivo* from autoradiography images. Various tumor cell density and cell spatial distribution measured on

pathology slides can be distinguished using TI calculated from autoradiography images despite the difference in voxel size and spatial resolution. Yet, tumor texture as assessed from PET images only coarsely reflects the spatial distribution and density of tumor cells.

Keywords: texture analysis; tumor heterogeneity; PET; autoradiography; histology

Running title: multi-scale texture analysis

INTRODUCTION

A current challenge in oncology is to offer each patient a personalized treatment accounting at best for the characteristics of the tumoral lesions. To achieve this, the specific histological or genetic features of the patient tumors have to be determined, including the intratumoral heterogeneity, before prescribing a therapy (*1*). This is most often performed based on the analysis of only a biopsy. Yet, a recent study (*2*) reported that 63 to 69% of all mutations were not detectable within a single tumor sample. Indeed, intratumoral heterogeneity is reflected by different mutations in sub-populations of cells distributed in different regions of the tumor, some of which being far from the biopsy site. Positron Emission Tomography (PET) provides molecular information regarding the whole tumor and might therefore be a relevant option to complete the characterization of the tumor. Indeed, PET imaging is non-invasive, can be repeated during the course of therapy and allows for a comprehensive assessment of each tumor site. Using different tracers, PET could even yield information regarding the glucose metabolism of cells (18F-FDG), tumor hypoxia (18F-FAZA, 18F-FMISO) or tumor cell proliferation (18F-FLT) (*3*). Recently, several studies have used tumor texture analysis in PET images to characterize intratumoral heterogeneity. Most reports assessed the ability of texture indices (TI) to predict outcome in patients with cancer and some have reported a relationship between these parameters and the tumor characteristics such as the Tumor-stage in esophageal cancers (*4*) or the tumor grade in glioma (*5*). Gao et al (*6*) showed that a combination of TI measured from PET and Computerized Tomography (CT) images before treatment could contribute to the assessment of the Node-stage. In the

context of the development of a computer-aided diagnosis system, Lartizien et al (*7*) demonstrated that using the 12 most discriminant PET and CT features, it was possible to discriminate between hypermetabolic cancer lesions and hypermetabolic inflammatory or physiological regions in patients with lymphoma, with an Area Under the Receiver Operating Characteristic curve equals to 0.91. To separate malignant from benign bone and soft-tissue lesions, Xu et al (*8*) used a combination of 4 TI calculated from PET images (entropy and coarseness) and CT images (entropy and correlation) and improved the lesion classification compared with the Standardized Uptake Value (SUV) only. In breast cancer, the combination of High Gray-level Run Emphasis and SUVmax identified triple-negative breast cancer lesions (*9*) with a sensitivity of 77% and a specificity of 71%. Despite these encouraging results, the biological interpretation of TI derived from PET and whether they reflect the microscopic intratumoral heterogeneity have not been investigated.

In this context, the purpose of this study was to understand the relationship between TI measured in 18F-FDG PET images, autoradiography images and pathological sections. The motivation was to determine whether TI measured at the PET scale (mm) was representative of features seen at a microscopic scale (autoradiography images) and at the cellular level (pathological slides).

MATERIALS AND METHODS

Animal models

All animal procedures were approved by the ethics committee in charge of animal experimentation (CETEA DSV n°44, reference 12-036) and were performed in accordance with European guidelines on handling laboratory animals.

Three mice bearing orthopically implanted mammary tumors derived from transgenic MMTV-PyMT mouse were used. Tumor donors were FVB/N-Tg (MMTV-PyMT) 634Mul/J (PyMT) 12-weeks-old mice. Aseptically collected mammary tumors from PyMT mice were minced and immersed in cold Dulbecco's Modified Eagle's Medium (Sigma, USA). Mechanical cell dissociation was performed using Medicon disposable chambers (BD bioscience, USA). The cell suspension was then progressively filtered using Filcon filters with pore sizes of 500 µm, 200 µm and 70 µm (BD bioscience). Finally, cells were aliquoted in freezing medium (Life Technologies, USA) and stored in liquid nitrogen. After freezing medium removal and enumeration, tumor cells were directly inoculated, without any in vitro culture step, in the mammary fat pad of the posterior nipple in FVB mice.

PET/CT imaging protocol

Thirty to thirty-six days after the implantation of tumors, the animals fasted for 12 hours before the PET/CT scan and they were anesthetized using 2% isoflurane. PET/CT was performed using the dedicated small animal Siemens Inveon PET/CT scanner. Each mouse was injected with 7.9 ± 0.3 MBq of 18F-FDG and the acquisition started 65 ± 5 min after this injection and lasted for 30 min (2 frames of 15-min acquisitions each). The energy window was set from 350 to 650 keV. PET

images were reconstructed using 2D ordered subset expectation maximization algorithm using 4 iterations and 16 subsets. Scatter correction was performed using the direct calculation from analytical formulas (*10*) and attenuation correction was based on the CT images. The PET images were not post-filtered. The matrix size was 256x256x159 voxels, corresponding to a PET voxel size of 388x388x796 μm^3. The PET volume was converted in SUV units normalized by mouse body weight (Fig. 1).

Tumor tissue sectioning

After the PET/CT acquisitions, all animals were sacrificed and the tumors were removed using landmarks regarding the orientation of the tumor in the mouse to facilitate the registration of autoradiography images and histological slides with PET images. A cryostat (Leica) was used to slice 20 µm-thick transaxial tissue sections from the resulting tissue block. Sections were obtained at 100 µm-intervals. Each section was put onto a glass slide. The first tumor yielded 116 slices while the other two led to 102 slices each.

Autoradiography imaging protocol

The tumor sections were exposed to imaging plates for 15 hours. The plates were scanned using a Storm scanner (GE Healthcare). The voxel size of autoradiography images was 50x50x20 μm^3 (Fig. 1).

Histochemistry

To investigate the biological meaning of TI, we assumed that FDG uptake was

correlated with the cellular density in tumors. The slices were thus analyzed using hematoxylin/eosin staining. Hematoxylin colors nuclei in blue-purple, while the cytoplasm and extracellular matrix are in pink (Fig. 1). To separate the hematoxylin component, we used the Color Deconvolution function (*11*) of the ImageJ software. The resulting "hematoxylin" images were resampled to match the autoradiography voxel size (50x50x20 μm^3) using TransformJScale function (*12*) from ImageJ, yielding resampled "hematoxylin" images called histological images thereafter (Fig. 1). From these images, a Volume of Interest (VOI-H) was segmented manually to separate the tumor from the background.

Creation and registration of 3D volumes

For each tumor, a 3D autoradiography volume was created from the autoradiography slices using the StrackReg function (*13*) of ImageJ. The 80 µm-interval between two slices were filled using a cubic-β spline interpolation (TransformJScale function of ImageJ) hypothesizing that morphological changes over this short distance were negligible. Three 3D autoradiography volumes were created (Fig. 1): the original autoradiography volume (voxel=50x50x20 μm^3) called VOI-AR, the resampled autoradiography volume (RAR) corresponding to the autoradiography volume resampled to the PET voxel size using the TransformJScale ImageJ function (voxel=388x388x796 μm^3), and the resampled (voxel=388x388x796 μm^3) and smoothed autoradiography volume (SRAR) corresponding to the RAR volume smoothed with a Gaussian filter (SD=677µm) to match the PET image resolution. The PET tumor volume was also registered to the SRAR volume using the so-called "Optimized Automatic Image Registration" of the Medical Image Processing

Analysis and Visualization (Mipav) software (*14*). This registration approach determines the transformation (12 degrees of freedom) that minimizes the correlation ratio function and involves a trilinear interpolation.

Tumor volumes of interest, VOI-PET, VOI-AR, VOI-RAR and VOI-SRAR, were segmented using a threshold equal to 40% of the maximum voxel intensity in the volume.

The autoradiography, RAR and SRAR volumes were normalized using:

$$I(SUV) = \frac{Weight(g) \times K^*(kBq) \times I(au)}{Dose(t=t_{auto})} \qquad (1)$$

$$K^*(kBq) = \frac{C_{PET}\left(\frac{kBq}{ml}\right) \times V_{PET}\,(ml)}{C_{auto}\left(\frac{au}{ml}\right) \times V_{auto}(ml)} \qquad (2)$$

where I(au) and I(SUV) correspond to the voxel value before and after normalization, respectively. Dose is the injected dose corrected for the radioactivity decay at the time of autoradiography. C_{PET} is the mean intensity in VOI-PET and V_{PET} is the volume of VOI-PET. Similarly, C_{auto} is the mean intensity in VOI-AR (or VOI-RAR or VOI-SRAR) and V_{auto} is the volume of VOI-AR (or VOI-RAR or VOI-SRAR). K* therefore represents the scaling factor that converts autoradiography images expressed in arbitrary units into SUV units so that they can be compared with the PET images.

Similarly, the histological images were normalized using a calibration factor K**:

$$I(SUV) = K^{**} \times I(au) \qquad (3)$$

$$K^{**}(SUV) = \frac{C_{auto}\left(\frac{SUV}{ml}\right)\times V_{auto}\,(ml)}{C_H\left(\frac{au}{ml}\right)\times V_H(ml)} \qquad (4)$$

C_H represents the mean intensity in VOI-H and V_H is the volume of VOI-H.

Texture analysis

For each selected 2D-image and each modality, we computed the maximum intensity in the tumor region of interest. Before computing the texture matrices, the voxel intensities were resampled using an absolute resampling method with fixed bounds and 64 discrete values (*15*):

$$R(x) = round[64 \times \frac{I(x) - Lower\ Bound}{Upper\ Bound - Lower\ Bound}] \qquad (5)$$

where R(x) corresponds to the voxel intensity after discretization and I(x) was the voxel intensity before discretization. The lower bound was set to 0. The upper bound corresponds to the maximum intensity over all images for each image type across the three tumors. The upper bound was 6 for the PET images, 0.1 for autoradiography images, 12 for RAR images, 9 for SRAR images and 175 for histological images.

After this resampling step, we computed three texture matrices in 2 dimensions for each slice and each modality and we extracted 6 TI with the LIFEx software (http://www.lifexsoft.org/): homogeneity and entropy computed from the co-occurrence matrix, Short-Run Emphasis (SRE) and Long-Run Emphasis (LRE) from the gray-level run length matrix and Low Gray-level Zone Emphasis (LGZE) and High Gray-level Zone Emphasis (HGZE) from the gray-level zone length matrix. These TI were selected given their previously demonstrated robustness with respect to the segmentation method (*16*).

Relationship between TI and actual biological tissue composition

Using a tumor slice including a great variety of cell density, we investigated the sensitivity of TI to various patterns of cell arrangements. To do so, we defined 80 sub-regions of 30 x 30 voxels in the histological image and copied them in the autoradiography image. The histological images were transformed into binary images using a threshold to identify the presence of cells. Sub-regions exhibiting three cell arrangements were identified: A: a large continuous region of cells with no more than three small islands (<15% of the surface) of extracellular matrix, or the opposite (mostly extracellular matrix with ≤3 small islands of cells); B: two distinct tissue types (cells and extracellular matrix) clearly separated in terms of location and in comparable proportions (between 35 and 65% of cells) and C: two nested mixed tissue types in comparable proportions (between 35 and 65% of cells). The 15%, 35% and 65% values were arbitrarily chosen to obtain well distinct cell arrangements. In each selected sub-region, we computed the value of TI in autoradiography images.

Statistical analysis

To investigate the relationship between TI measured in the various image types, we plotted 1) TI measured in the PET images as a function of TI measured in the autoradiography images; 2) TI measured in the autoradiography images as a function of TI measured in the histological images; 3) TI measured in the PET images as a function of TI measured in the histological images. All these plots were

characterized by the Spearman correlation coefficient (R). Moreover, we compared the TI values measured in PET and autoradiography images using Bland-Altman plots.

We studied the impact of voxel size and spatial resolution by plotting: 1) TI from autoradiography against TI from RAR; 2) TI from RAR against TI from SRAR, and also characterized the observed relationships using Spearman correlation coefficients.

We studied the ability of TI to distinguish between the three types of sub-regions defined in the histological images based on box-plot analyses and Wilcoxon tests. Similarly, we studied the impact of cell density in the A sub-regions defined in the histological images by comparing TI calculated in A sub-regions with a majority of cells (percentage of cells higher than 85%, denoted A+ thereafter) and TI calculated in the A sub-regions with a majority of extracellular matrix (percentage of cells lower than 15%, denoted A- thereafter). Again, the 15% and 85% thresholds were arbitrarily chosen.

RESULTS

We selected 28 slices from the three tumors with a minimum interval equal to 388 µm between two slices in PET, autoradiography, RAR, SRAR and histological volumes.

Correlation between texture analysis at different scales

The values of TI extracted from autoradiography images are plotted against the TI measured from PET images in Figure 2 and Supplemental Figure 1. All TI and maximum intensity were significantly correlated between PET and autoradiography images with correlation coefficient between 0.57 and 0.85 (Table 1). The results also show that the absolute values of TI and maximum intensity differ between the two modalities according to the Bland-Altman plots (Fig. 2; Supplemental Fig. 1). For example, entropy, SRE, HGZE and maximum intensity computed from autoradiography images were lower than the same indices extracted from PET images. Inversely, homogeneity, LRE and LGZE calculated from autoradiography images were higher than these indices computed from PET images.

TI extracted from pathological images are plotted against TI computed from autoradiography or PET images in Figure 3 and Supplemental Figure 2. Neither the TI nor the maximum intensity were significantly positively correlated between histological and autoradiography or PET images (Table 1).

Impact of voxel size and spatial resolution in autoradiography images

The plots of the TI values measured in the RAR images against those measured in the autoradiography images are shown in Supplemental Figure 3, together with the values measured in the SRAR images against the values measured in the RAR images. The Spearman correlation coefficient characterizing the relationship between RAR and autoradiography or SRAR and RAR are given in Table 1. All TI and the maximum intensity were moderately or strongly correlated between autoradiography and RAR ($R \geq 0.62$) but the absolute value of TI differed. Indeed, the texture measured from autoradiography images was more homogeneous than that measured from RAR images. For instance (Supplemental Fig. 3), homogeneity, LRE and LGZE were greater in the autoradiography images than in the RAR images, and the opposite was true for the other indices, demonstrating that the voxel size influences the absolute value of TI.

Regarding the influence of spatial resolution, the plots of TI measured from SRAR images as a function of TI measured in RAR images (Supplemental Fig. 3) showed a high correlation for entropy ($R=0.96$), while the correlation was only moderate for the other TI and for the maximum intensity ($R=[0.59\text{-}0.69]$). For homogeneity, entropy, SRE and LRE, there was no significant difference between the values calculated in the RAR and SRAR images (pvalue of Wilcoxon test > 5%), suggesting that changing the spatial resolution did not change significantly the TI values when the voxel size is identical.

TI values in autoradiography images as a function tumor cell density and arrangement

In one slice selected for the great diversity of tumor cell percentage and arrangement (Fig. 4), we computed TI from the AR image in 30 out of the 80 sub-regions of 30x30 voxels. These 30 sub-regions were selected so that they clearly exhibited one the three types of cell pattern. Only homogeneity, entropy, SRE and LRE were significantly different between the three cell patterns (Fig. 5; Supplemental Fig. 4): the A sub-regions were more homogeneous than C that were in turn more homogeneous than B. For instance, homogeneity and entropy had a median value of 0.68 and 1.42 in the A regions, 0.61 and 1.96 in B and 0.64 and 1.81 in C. Conversely, HGZE, LGZE and maximum intensity were not significantly different between the three cell arrangements.

Among the A sub-regions, HGZE, LGZE and maximum intensity could separate sub-regions with a majority of cells (>85% of cells, A+) from sub-regions composed of a majority of extracellular matrix (<15% of cells, A-), while other indices could not. These TI were sensitive to the cell density. For example, HGZE and maximum intensity were higher in A+ than in A- (Fig. 5; Supplemental Fig. 4), whereas LGZE varied in an opposite direction (Supplemental Fig. 4). In A+ sub-regions, the median value was 965 for HGZE, 0.06 for maximum intensity and 0.001 for LGZE against respectively 163, 0.03 and 0.007 in A- sub-regions.

DISCUSSION

In this study, we showed that the texture computed from *in vivo* PET images reflects the texture computed from *ex vivo* autoradiography images and we found that some TI were sensitive to the spatial distribution of cells while others were mostly sensitive to the density of tumor cells in the region of interest. This study also demonstrates the potential and limitations of texture analysis from PET images to quantify the microscopic heterogeneity.

Multi-scale comparison of TI values

All TI were significantly correlated between PET and autoradiography images, even if the absolute values differ between the two modalities due to the difference in voxel size and spatial resolution. This demonstrates the ability of TI to quantify heterogeneity seen at two different scales. However, we observed that TI computed from PET or autoradiography images were not significantly correlated with TI computed from hematoxylin images, suggesting that the precise cell spatial distribution was neither captured by PET nor by autoradiography images. These results might be partly explained by the non-specific uptake of FDG. Indeed, cells seen on the pathological slides were not necessarily FDG-avid, and a better correlation might be expected using GLUT-1 immunohistochemistry (*17*). Nevertheless, using sub-regions in autoradiography images, we demonstrated that some TI were sensitive to the spatial organization of tumor cells while others were sensitive to the cell density as seen in the hematoxylin images. TI may therefore reflect some useful information regarding the spatial organization of tumor cells even if a TI value measured from a PET or autoradiography image cannot be directly

translated into a textural pattern seen in the pathological slides. This could explain some results previously reported in the literature. For instance, texture analysis from Single-Photon Emission Computed Tomography (SPECT) images (*18*) could differentiate between two hepatic metastatic colorectal cancers with a greater homogeneity for the well-differentiated than poorly-differentiated tumors. They also found that, in poorly-differentiated liver metastases, TI were sensitive to the action of an anti-vascular treatment that produced an increase of homogeneity. Inversely, TI could not separate between treated and un-treated well-differentiated metastases. This example illustrates the limited characterization of tumor cell heterogeneity using TI calculated from SPECT images as histological analysis exhibited the modifications induced by the treatment in the metastases.

Robustness of TI with respect to voxel size and image spatial resolution

Our results show that all TI were significantly correlated between autoradiography and RAR or RAR and SRAR with a correlation coefficient higher than 0.59, suggesting that TI measured from PET images reflect TI measured at a microscopic scale. Yet, we observed that the absolute value of all TI was affected by substantial differences in voxel size and also for some TI by the spatial resolution. As a result, it appears that TI values cannot be readily compared between images with large difference in voxel size (PET voxel size was 2,000 times larger than autoradiography images). Our results show that autoradiography images are seen by TI as more homogenous than PET images, which is consistent with a higher homogeneous texture in autoradiography images than in RAR images. Although this might appear as counter-intuitive, this trend is due to the calculation procedure. When calculating

TI, the actual voxel size is not taken into account. Yet, the autoradiography images contain more voxels than PET or RAR images for the same lesion. As a result, in autoradiography images, a small homogeneous zone, which might be represented by only very few voxels in the PET images, involves many voxels hence “weights” more and makes the TI reflect a more homogeneous texture than in the PET image. Entropy was the most robust with respect to differences in spatial resolution with R=0.96 between RAR and SRAR images with similar values between the two images (between 1.8 and 2.8). These results were consistent within the correlation groups previously identified (*16*), as reported in Supplemental Table 1. In (*19*), it was shown that the 6 TI were moderately influenced by the smoothing of PET images performed post-reconstruction except LGZE. These results were obtained with relative resampling consisting of discretizing intensities between the minimum and maximum intensities of each lesion, and not with the absolute resampling method used here. The smoothing was also weaker in (*19*) with a Full Width at Half Maximum equal to 0.9 times the voxel size compared with 1.7 times the voxel size between RAR and SRAR in our study. Consistent with our results, another study (*20*) demonstrated that the 6 TI we studied were equally or less impacted by the smoothing than SUVmax, also with the relative resampling. This is due to the resampling step performed before TI calculation. Indeed, by assigning the same value to voxels with close intensities, the resampling actually acts as a smoothing operation and reduces the impact of noise that affects other indices such as SUVmax.

A limitation of our study is the small number of animals and using only one tumor model. Moreover the uptake of 18F-FDG is not specific to tumor cells and can also

reflect inflammation. Still, this does not invalidate our results regarding the change in TI as a function of voxel size, spatial resolution, and observation scale (autoradiography or PET images). In addition, texture analysis was performed in 28 distant slices that could be seen as independent, so that correlations between TI calculated from the PET images and autoradiography were based on 28 points. Using more specific PET tracers, other immunohistochemical analysis, more animals and tumor cell lines would be necessary to better understand the relationship between heterogeneity seen on PET images and microscopic heterogeneity. In addition, the pre-clinical PET scanner has a spatial resolution of 1.6 mm and tumor diameters were approximatively equal to 10 mm. In a clinical setting, this would consist in analyzing tumors that are about 4 cm in diameter with a PET image spatial resolution of 6 mm. Further studies are needed to better understand the relationship between TI calculated from clinical images with biological tissue heterogeneity as seen on pathological specimens.

CONCLUSION

TI measured from PET images are well correlated with the corresponding TI calculated from autoradiography images, although the TI magnitudes are different due to different voxel size and different spatial resolution. Yet, TI computed from PET images cannot be easily correlated with texture measured from the pathological slides. Still, TI derived from autoradiography images are sensitive to the spatial distribution or density of tumor cells. Even if it does not capture precisely the heterogeneity of tumor cells, texture analysis might still provide addition information on lesions and therefore potentially on tumor aggressiveness and resistance of

therapy. Although the biological meaning of TI still needs to be clarified, accounting for TI to characterize a tumor could thus be useful to provide a more extensive characterization of the whole tumor compared to that obtained using a biopsy. Further investigations are needed to better elucidate the relationship between the tumor biological features and the heterogeneity of tracer uptake reflected by TI.

ACKNOWLEDGMENT

This work is supported by the "Lidex-PIM" project funded by the IDEX Paris-Saclay, ANR-11-IDEX-0003-02.

REFERENCES

1. Lianos GD, Mangano A, Rausei S, Katsios C, Roukos DH. Tumor heterogeneity-based resistance guides personalized cancer medicine. *Future Oncol*. 2014;10:1889–1892.

2. Gerlinger M, Rowan AJ, Horswell S, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med*. 2012;366:883–892.

3. Pantaleo MA, Nannini M, Maleddu A, et al. Conventional and novel PET tracers for imaging in oncology in the era of molecular therapy. *Cancer Treat Rev*. 2008;34:103–121.

4. Dong X, Xing L, Wu P, et al. Three-dimensional positron emission tomography image texture analysis of esophageal squamous cell carcinoma: relationship between tumor 18F-fluorodeoxyglucose uptake heterogeneity, maximum standardized uptake value, and tumor stage. *Nucl Med Commun*. 2013;34:40–46.

5. Pyka T, Gempt J, Hiob D, 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;43:133–141.

6. Gao X, Chu C, Li Y, et al. The method and efficacy of support vector machine classifiers based on texture features and multi-resolution histogram from 18F-FDG PET-CT images for the evaluation of mediastinal lymph nodes in patients with lung cancer. *Eur J Radiol*. 2015;84:312–317.

7. Lartizien C, Rogez M, Niaf E, Ricard F. Computer-aided staging of lymphoma patients with FDG PET/CT imaging based on textural information. *IEEE J Biomed Health Inform.* 2014;18:946–955.

8. Xu R, Kido S, Suga K, et al. Texture analysis on (18)F-FDG PET/CT images to differentiate malignant and benign bone and soft-tissue lesions. *Ann Nucl Med*. 2014;28:926-935.

9. Soussan M, Orlhac F, Boubaya M, et al. Relationship between tumor heterogeneity measured on FDG-PET/CT and pathological prognostic factors in invasive breast cancer. *Plos One*. 2014;9:e94017.

10. Watson CC. New, faster, image-based scatter correction for 3D PET. *IEEE Trans Nucl Sci.* 2000;47:1587–1594.

11. Ruifrok AC, Johnston DA. Quantification of histochemical staining by color deconvolution. *Anal Quant Cytol Histol.* 2001;23:291–299.

12. Meijering EH, Niessen WJ, Viergever MA. Quantitative evaluation of convolution-based methods for medical image interpolation. *Med Image Anal.* 2001;5:111–126.

13. Thevenaz P, Ruttimann UE, Unser M. A pyramid approach to subpixel registration based on intensity. *IEEE Trans Image Process.* 1998;7:27–41.

14. McAuliffe MJ, Lalonde FM, McGarry D, Gandler W, Csaky K, Trus BL. Medical Image Processing, Analysis and Visualization in clinical research. *In: 14th IEEE Symposium on Computer-Based Medical Systems*, *Proceedings.* 2001:381–386.

15. 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*. 2015;10:e0145063.

16. Orlhac F, Soussan M, Maisonobe J-A, Garcia CA, Vanderlinden B, Buvat I. Tumor texture analysis in 18F-FDG PET: relationships between texture parameters, histogram indices, standardized uptake values, metabolic volumes, and total lesion glycolysis. *J Nucl Med*. 2014;55:414–422.

17. Lee AF, Gown AM, Churg A. IMP3 and GLUT-1 immunohistochemistry for distinguishing benign from malignant mesothelial proliferations. *Am J Surg Pathol.* 2013;37:421–426.

18. Rajkumar V, Goh V, Siddique M, et al. Texture analysis of (125)I-A5B7 anti-CEA antibody SPECT differentiates metastatic colorectal cancer model phenotypes and anti-vascular therapy response. *Br J Cancer.* 2015;112:1882-1887.

19. Doumou G, Siddique M, Tsoumpas C, Goh V, Cook GJ. The precision of textural analysis in 18F-FDG-PET scans of oesophageal cancer. *Eur Radiol*. 2015;25:2805-2812.

20. Yan J, Chu-Shern JL, Loi HY, et al. Impact of Image Reconstruction Settings on Texture Features in 18F-FDG PET. *J Nucl Med*. 2015;56:1667–1673.

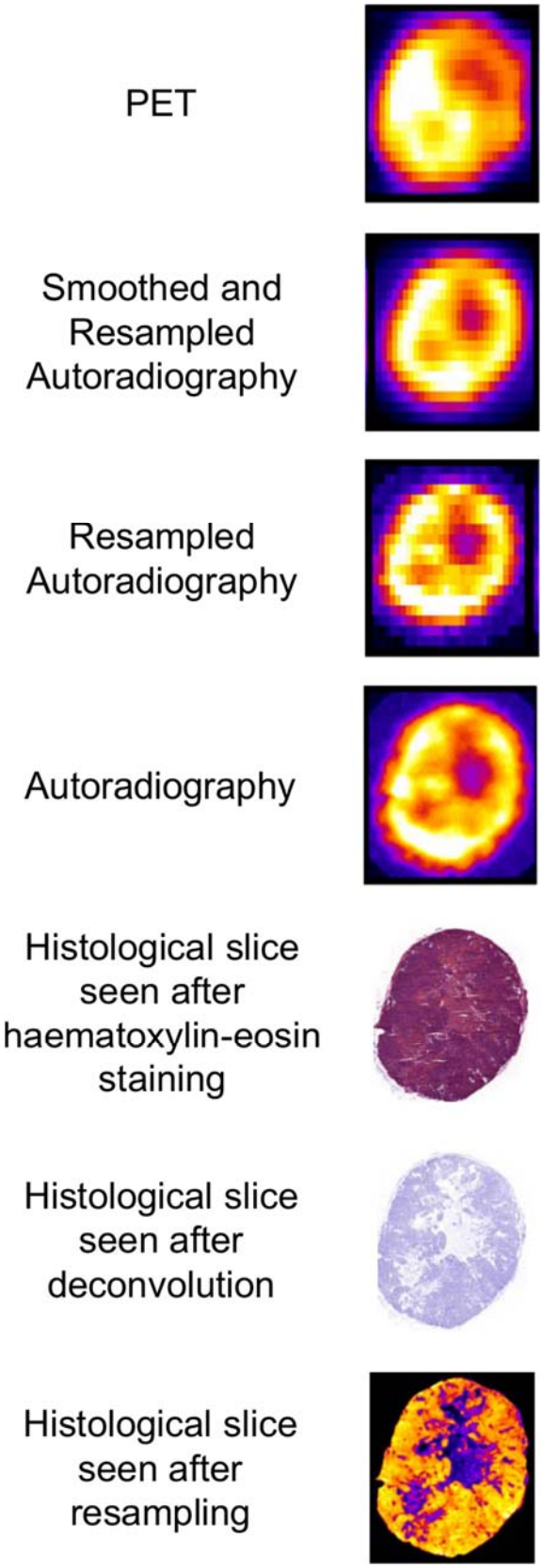

FIGURE 1: Example of tumor slice as seen in the different modalities.

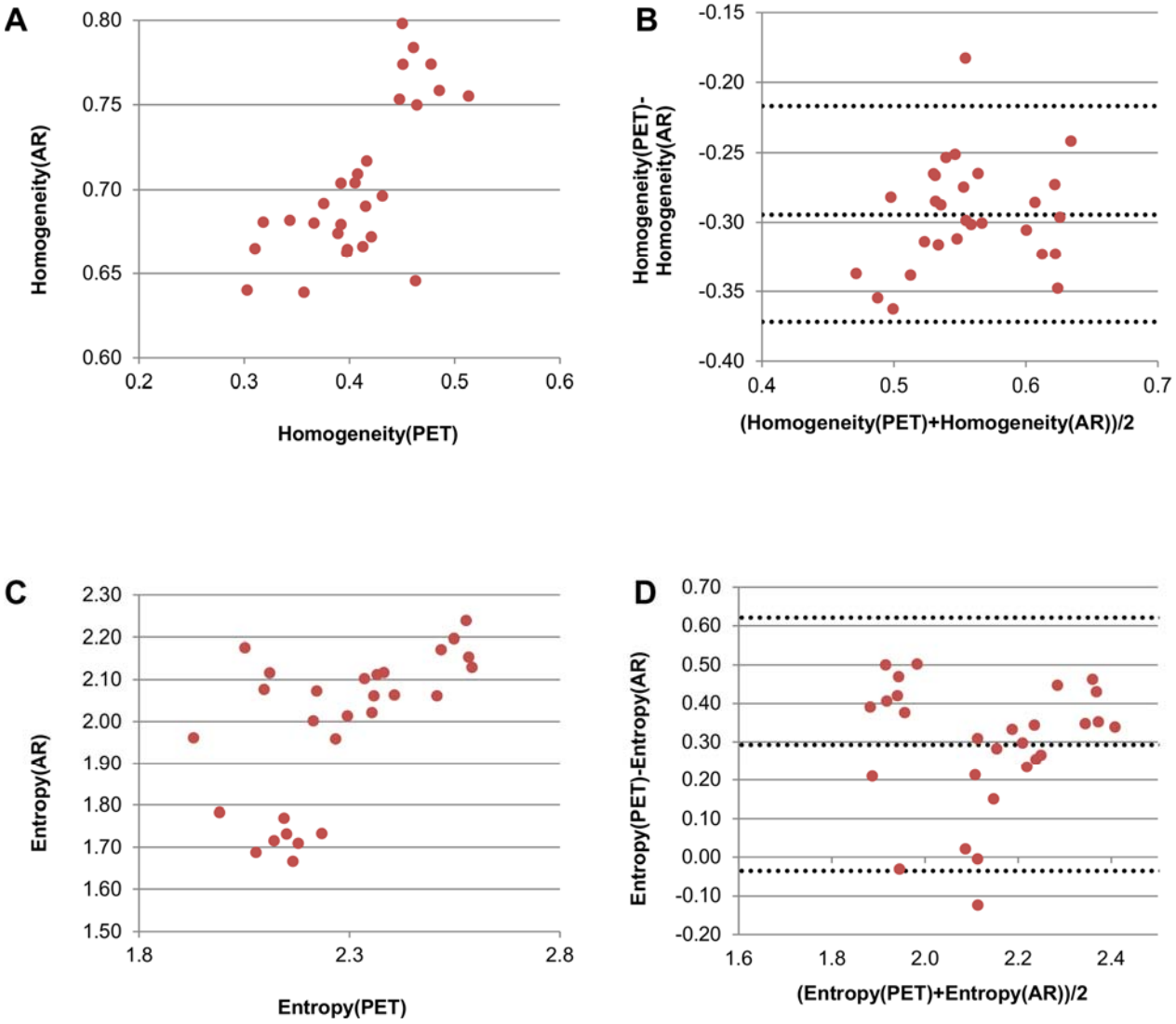

FIGURE 2: Plots of index measured from autoradiography (AR) images as a function of index measured from PET images and corresponding Bland-Altman plots.

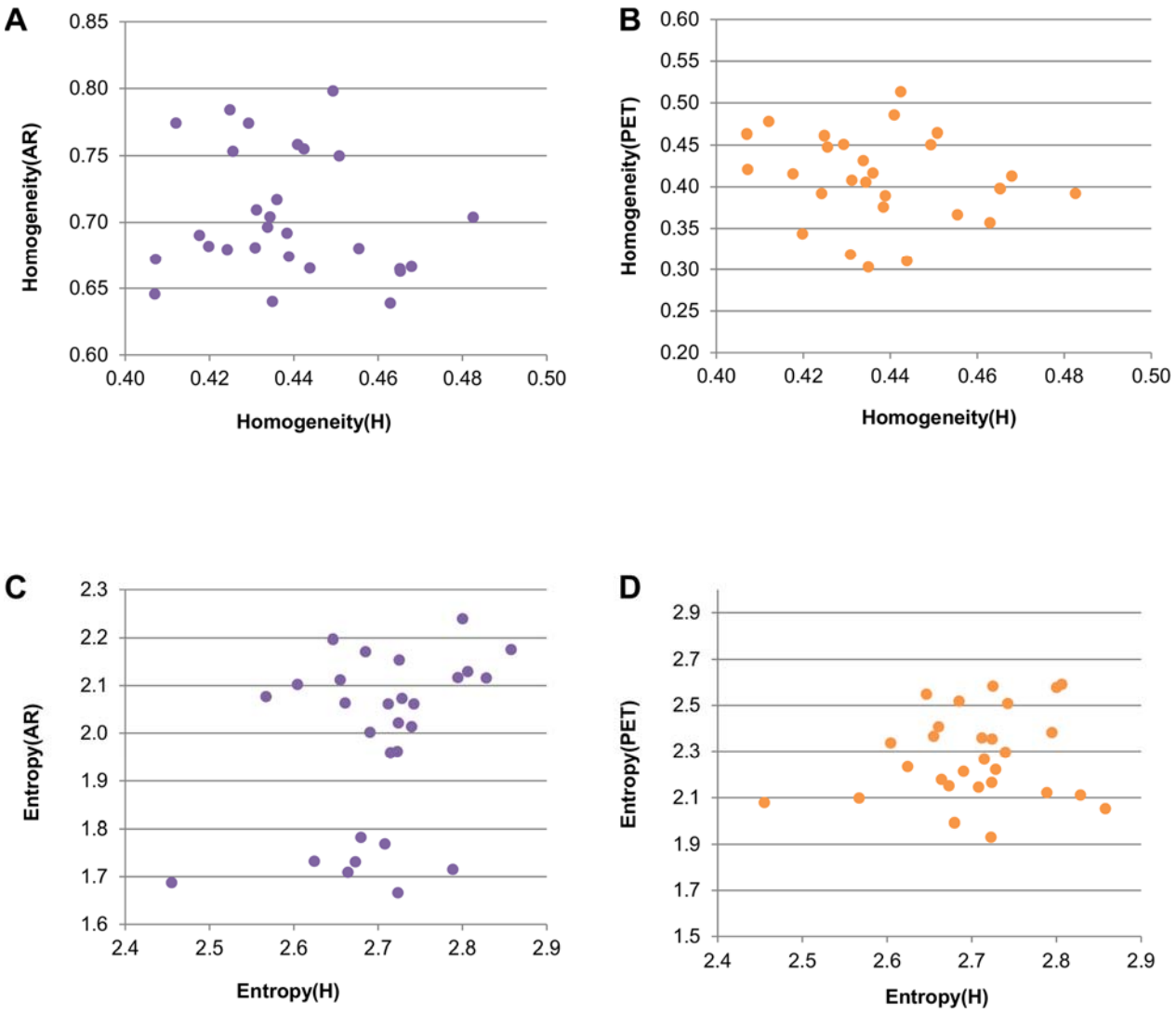

FIGURE 3: Plots of index measured from autoradiography (AR) images (in purple) or PET images (in orange) as a function of index measured from histological slides.

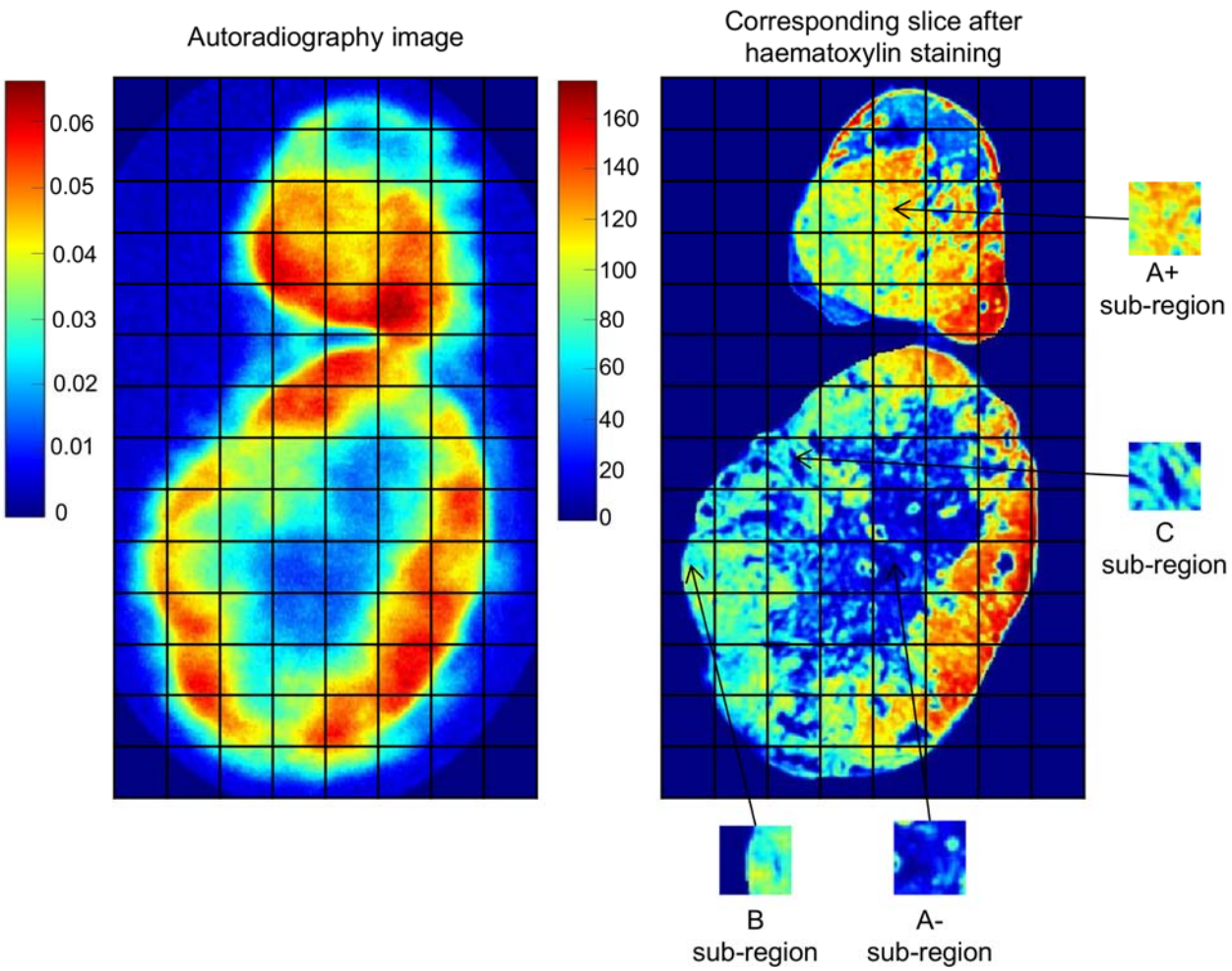

FIGURE 4: Image of selected tumor slice as seen in autoradiography and after haematoxylin staining.

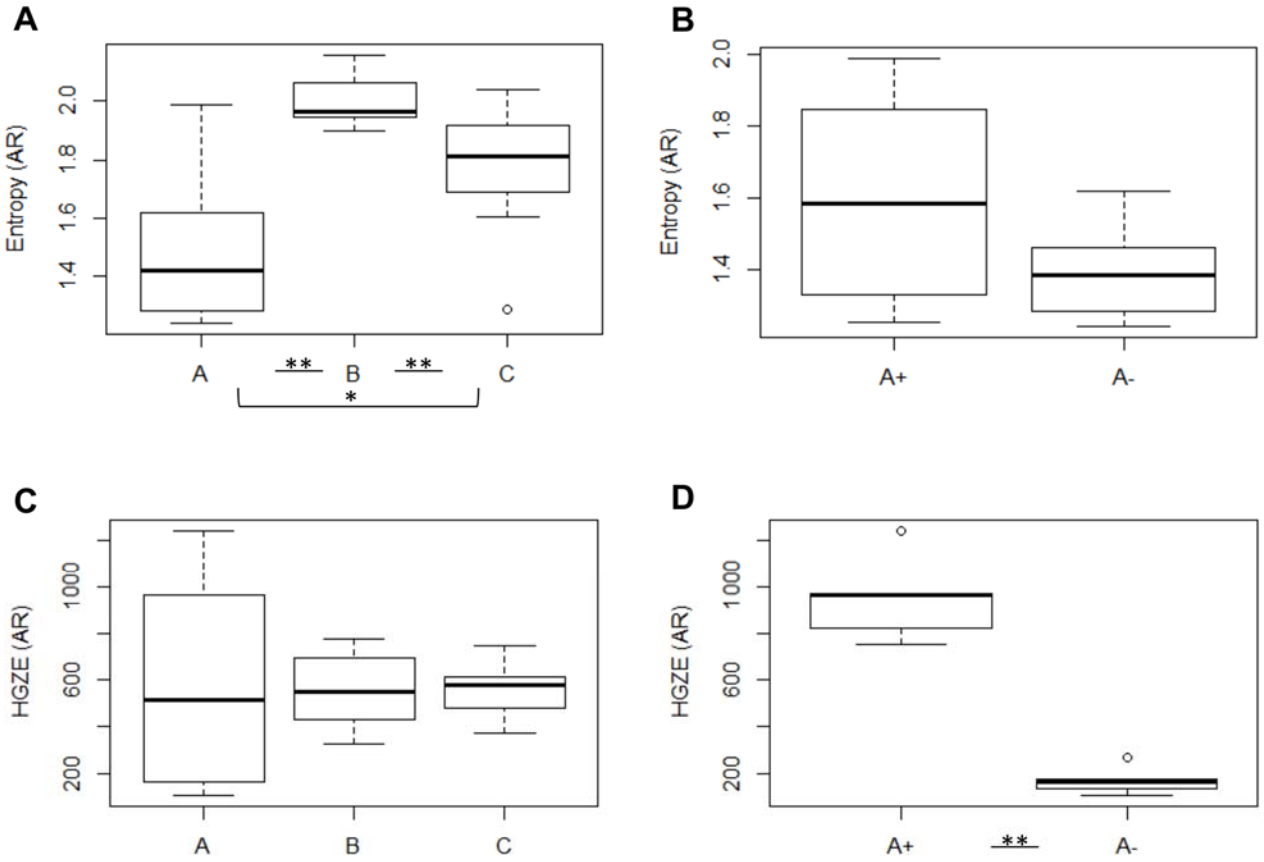

FIGURE 5: Boxplots of TI computed from autoradiography images as a function of the cell pattern (A, B and C) or as a function of the number of cells in homogeneous sub-regions (A+, A-). *: pvalue of Wilcoxon's test < 5%. **: pvalue of Wilcoxon's test < 1%.

TABLE LEGEND

TABLE 1: Spearman correlation coefficient (R) between different modalities for 6 TI and for the maximum intensity.

	PET vs. AR	H vs. AR	H vs. PET	AR vs. RAR	RAR vs. SRAR
Homogeneity	0.66*	-0.18	-0.23	0.71*	0.69*
Entropy	0.57*	0.31	0.13	0.77*	0.96*
SRE	0.67*	-0.25	-0.13	0.65*	0.60*
LRE	0.70*	-0.41*	-0.54*	0.62*	0.62*
LGZE	0.83*	-0.23	-0.06	0.92*	0.66*
HGZE	0.85*	0.32	0.29	0.78*	0.59*
Maximum intensity	0.75*	0.20	0.09	0.89*	0.68*

*: pvalue of Spearman correlation coefficient < 5%
AR: autoradiography images
H: histological images
RAR: resampled autoradiography images
SRAR: smoothed and resampled autoradiography images