HAL
open science

Can we trust the calculation of texture indices of CT images? A phantom study

Caroline Caramella, Adrien Allorant, Fanny Orlhac, Francois Bidault, Bernard Asselain, Samy Ammari, Patricia Jaranowski, Aurelie Moussier, Corinne Balleyguier, Nathalie Lassau, et al.

To cite this version:

Caroline Caramella, Adrien Allorant, Fanny Orlhac, Francois Bidault, Bernard Asselain, et al.. Can we trust the calculation of texture indices of CT images? A phantom study. Medical Physics, 2018, 45 (4), pp.1529 - 1536. 10.1002/mp.12809 . hal-03326446

HAL Id: hal-03326446

https://hal.science/hal-03326446v1

Submitted on 26 Aug 2021

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.

Medical Physics, 45(4), April 2018
doi.org/10.1002/mp.12809

Can we trust the calculation of texture indices of CT images? A phantomstudy

Caroline Caramella[a)]
Department of Imaging, Gustave Roussy Cancer Campus, Villejuif Cedex, France IR4M-UMR8081, CNRS, Univ Paris Sud, University Paris Saclay, Orsay Cedex, France

Adrien Allorant
Department of Biostatistics, Gustave Roussy Cancer, Villejuif, France

Fanny Orlhac
Imagerie Moléculaire In Vivo, INSERM CEA CNRS University Paris Sud University Paris Saclay CEA, Orsay, France INSERM U1030, Villejuif, France

Francois Bidault
Department of Imaging, Gustave Roussy Cancer Campus, Villejuif Cedex, France IR4M-UMR8081, CNRS, Univ Paris Sud, University Paris Saclay, Orsay Cedex, France

Bernard Asselain
IR4M-UMR8081, CNRS, Univ Paris Sud, University Paris Saclay, Orsay Cedex, FranceResearch Department, Gustave Roussy Cancer Campus, Villejuif, France

Samy Ammari, Patricia Jaranowski, and Aurelie Moussier
Department of Imaging, Gustave Roussy Cancer Campus, Villejuif Cedex, France

Corinne Balleyguier
Department of Imaging, Gustave Roussy Cancer Campus, Villejuif Cedex, France IR4M-UMR8081, CNRS, Univ Paris Sud, University Paris Saclay, Orsay Cedex, France

Nathalie Lassau
IR4M-UMR8081, CNRS, Univ Paris Sud, University Paris Saclay, Orsay Cedex, FranceResearch Department, Gustave Roussy Cancer Campus, Villejuif, France

Stephanie Pitre-Champagnat
IR4M-UMR8081, CNRS, Univ Paris Sud, University Paris Saclay, Orsay Cedex, France

(Received 12 June 2017; revised 17 January 2018; accepted for publication 26 January 2018; published 13 March 2018)

Purpose: Texture analysis is an emerging tool in the field of medical imaging analysis. However, many issues have been raised in terms of its use in assessing patient images and it is crucial to harmo-nize and standardize this new imaging measurement tool. This study was designed to evaluate the reliability of texture indices of CT images on a phantom including a reproducibility study, to assess the discriminatory capacity of indices potentially relevant in CT medical images and to determine their redundancy.
Methods: For the reproducibility and discriminatory analysis, eight identical CT acquisitions were performed on a phantom including one homogeneous insert and two close heterogeneous inserts. Texture indices were selected for their high reproducibility and capability of discriminating different textures. For the redundancy analysis, 39 acquisitions of the same phantom were performed usingvarying acquisition parameters and a correlation matrix was used to explore the 2 9 2 relationships.LIFEx software was used to explore 34 different parameters including first order and texture indices. Results: Only eight indices of 34 exhibited high reproducibility and discriminated textures fromeach other. Skewness and kurtosis from histogram were independent from the six other indices but were intercorrelated, the other six indices correlated in diverse degrees (entropy, dissimilarity, and contrast of the co-occurrence matrix, contrast of the Neighborhood Gray

Level difference matrix, SZE, ZLNU of the Gray-Level Size Zone Matrix).
Conclusions: Care should be taken when using texture analysis as a tool to characterize CT images because changes in quantitation may be primarily due to internal variability rather than from real phy-sio-pathological effects. Some textural indices appear to be sufficiently reliable and capable to dis- criminate close textures on CT images. *© 2018 American Association of Physicists in Medicine* [https://doi.org/10.1002/mp.12809]

Key words: computed tomography, phantom study, texture analysis

1. INTRODUCTION

Texture is defined as the quantitation of the spatial distribution of repeating patterns. It is used to translate the homogeneous or nonhomogenous appearance of the surface of an object on an image. It is also defined as a two-dimensional phenomenon: the first dimension is a description of the basic elements, or the "primitive" (the pattern), from which the texture is formed; the second dimension relates to the description of the spatial organization of these primitives.[1] Radiomics is a relatively new discipline for the clinical integration of the quantitative features of digital medicalimages as determined through mathematical analysis. One potential quantitative image feature of clinical significance istumor tissue heterogeneity. Many studies have already explored the contribution of texture analysis in the field of oncology, addressing a range of topics (diagnosis, prognosis,correlation with histological or biological characteristics of the tumors) with wide-ranging results.[2,3] The main difficulty associated with this area is that there are currently as many ways of calculating texture indices as there are researchgroups focusing on texture analysis.

Several issues have already been raised: the definition of texture algorithms, which texture indices are relevant in medical images and the lack of reproducibility of texture calculations[4–7] It is crucial to the success of this method- ology to harmonize the procedures and textural features and to identify any inadequate parameters which could jeopardize the chances of obtaining useful results in clinical practice.

Our study is a first step toward the standardization of this new imaging measurement tool. It was designed to evaluate the reliability of texture indices of CT images on a phantom including a reproducibility study, to assess the discriminatory capacity of indices potentially relevant in CT medical images, and to determine their redundancy.

2. MATERIALS AND METHODS

2.A. Phantom

We built a phantom based on the commercially available CIRS® Electron Density Phantom Model 062M (Norfolk, VA, USA) which allows adequate maneuverability and the option of filling 17 holes with commercial homogeneous plugs as well as with "home-made" inserts.

The cylindrical phantom dimensions are 180 mm in diameter and 50 mm in thickness. Plug dimension is 30 mm in diameter and 50 mm in thickness. A homogenous plug with the same density as that of muscle (physical density of 1.06 g/cc) was used (named C1). In order to mimic human tissue, we designed two additional heterogeneous mixtures based on Ecoflex® (BASF, Ludwigshafen, Germany) which is a stable polymer, composed of two pure carbon fragmentsin different proportions randomly arranged in an agarose support (named C2 and C3), these two inserts have the advantageof having a similar range of physical density and histograms as those found in human soft tissue, such as tumoral lung tis-sues. Figure 1 shows the inserts and the corresponding image of a CT acquisition slice.

2.B. Reproducibility study

Eight identical CT acquisitions were made with a Discovery CT750 HD with 64 multidetector arrays (GE Healthcare,Milwaukee, WI, USA) using parameters routinely used for standard thoracic CT acquisition in patients treated in our center: voltage 120 kVp; x-ray tube current 150 mAs; 1.25 mm slice thickness; 1.375 pitch; rotation time 0.6 s, field of view 360 mm; reconstruction algorithm ASIR 0%; standard filter. Images were acquired sequentially and the phantom setup was not altered.

2.C. Correlation of the texture index

For assessment of redundant indices, we performed 39 supplementary CT acquisitions on the same phantom andthe same CT device, which differed for the acquisition parameters [voltage, intensity, slice thickness, pitch,

rota- tion time, field of view, reconstruction algorithm, and fil-ter (Table I)].

2.D. Texture analysis

Index extraction was performed with LIFEx (Local Image Features Extraction) open access software (http://www.lifex soft.org; Orsay, France)[8] which calculates the minimum and maximum density values, as well as the mean and standard deviation of the gray values of the region of interest (ROI). From these primary calculations, the software accesses histogram values (skewness, kurtosis, entropyH, energyH), the co-occurrence matrix values (homogeneity, energy, contrast,correlation, entropy, dissimilarity), the Gray-Level Run Length Matrix GLRLM (SRE/LRE, LGRE/HGRE, SRLGE/SRHGE, LRLGE/LRHGE, GLNU/RLNU, RP), the Neigh-
borhood Gray-Level difference matrix NGLDM (coarseness,busyness, contrast), and the Gray-Level Size Zone Matrix GLSZM (SZE/LZE, LGZE/HGZE, SZLGE/SZHGE,LZLGE/LZHGE, GLNUz/ZLNU, ZP).
The description of texture indices is provided in Data S1. An additional absolute discretization pixel intensity in Hounsfield units (HU) was applied to promote differentiationof intensities before texture index extraction.[9] We chose 200gray values, by steps of 10 HU, between the extreme values of 1000 and +1000 HU. Two-dimensional circular ROI were drawn on each of the three inserts. The size of these ROI was identical for each insert (458 mm², 758 pixels) (Fig. 2).

2.E. Statistical analyses

The eight CT acquisitions resulted in eight tables of 34 texture indices from the three ROI. The reproducibility of thetexture analysis was estimated from the coefficient of variation (CV). CVs were calculated for each texture index according to the following formula:

$$cv = \frac{\sigma}{\mu}$$

where r is the standard deviation and l the mean value of the data sample.

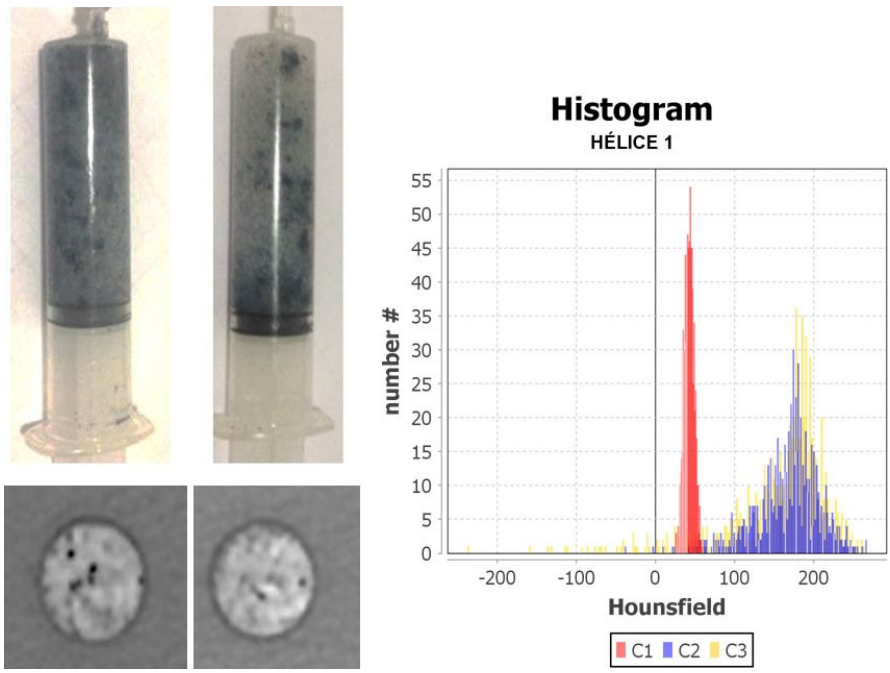

FIG. 1. C2 and C3 inserts compacted into two syringes adapted to precisely fill the holes of the CIRS® 062M phantom. Resulting histogram of the three ROI.[Color figure can be viewed at wileyonlinelibrary.com]

Variation in CV relating to imaging biomarkers is not standard,[10] in this study, a CV < 20% was considered acceptable.

Dot plots were performed to show the dispersion of valuesof the different indices among C1, C2, and C3. A nonparametric Mann–Whitney test was used to exclude the indices which could not differentiate C1 from C2/C3, and C2 from C3. Indices that would be unable to distinguish a homogeneous area from a heterogeneous area were considered to be

of no interest in the assessment of tumoral heterogeneity, as were indices that would not be able to distinguish two "visible to the naked eye" heterogeneous areas.

A Pearson correlation between the texture indices was

assessed to evaluate their redundancy on the 39 different CTacquisitions. A correlation matrix was used to explore the 2 9 2 relationships between the retained parameters. Principal component analysis was used to represent the different retained

indices and their pattern of correlation on a simple graph. Principal component analysis is a multivariate technique that analyzes a data table in which observations are described by several intercorrelated quantitative dependent variables. Its goal is to extract the important information from the table, torepresent it as a set of new orthogonal variables called principal components, and to display the pattern of similarity of theobservations and of the variables as points in maps.[11]

3. RESULTS

The characteristics of intensities (in HU) obtained within each ROI (C1, C2, and C3) are reported in Table II. As expected, the homogeneous ROI C1 exhibits a very narrow distribution spectrum of the intensity values of the pixels. The maximal variability was observed on intensities for the minValues at 16% and decreased to 4% for maxValues and meanValues.

Heterogeneous ROIs C2 and C3 exhibited a close histogram distribution of intensity values of pixels, with a range close to those of human tissues. Mean values of C2 and C3 were close (<6%) despite differences in the minimal and standard pixel values. The maximal variabilities were also observed for the minValues at 19% and were less than 1% forthe other values. The fluctuation of the mean value of C1 is larger than in C2 and C3, this may be explained by the fact that the stochastic noise effect may be more marked in a homogeneous area than in heterogeneous areas.

The CVs of texture indices for the three ROI pooled fromthe eight identical CT acquisitions are shown in Table III. Values obtained from the homogeneous insert C1 were morevariable than the C2 and C3 inserts, with CVs ranging from 1% to 26%. Of note, the high variations of skewness and busyness in CV values could be explained by the variations around 0 of the mean l values of these indices. In the case of busyness, the strong variability implied a lack of reliability. The maximum variations observed for C2 and C3 were between 0% and 18%.

Figures 3, 4, and 5 were built to assess which indices per-mitted discrimination of each insert. The three indices deter-mined as unable to discriminate C1 from C2 and C3 were busyness, GLNU from the Gray-Level Size Zone Matrix GLSZM (Fig. 3), and SZLGE. The following parameters were identified as unable to discriminate C2 from C3: entropy and energy derived from histogram; homogeneity, energy, and correlation (Fig. 4) from the co-occurrence matrix; SRE, LRE, LGRE, HGRE, SRLGE, SRHGE, LRLGE, LRHGE, RLNU, RP from the Gray-Level Run Length Matrix; and SZE, LZE, LGZE, HGZE, SZHGE, LZLGE, LZHGE, GLNU, ZP from the Gray-Level Size Zone Matrix.

From the repeatability study, eight indices were selected because they allow the discrimination of the three textures between them. These indices are skewness (Fig. 5), kurtosis derived from histogram, contrast, and entropy from the co- occurrence matrix, dissimilarity and contrast from the Neighborhood Gray-Level difference matrix, SZE and ZLNU from the Gray-Level Size Zone Matrix.

All eight indices were associated with highly significant P-values (<0.0002; Mann–Whitney test) when comparing C1 to C2, C2 to C3, and C1 to C3, as there was no overlap between the distributions for any of the eight parameters.

TABLE I. Design of the 40 different CT acquisitions performed for the redundancy analysis. The first CT acquisition corresponds to the acquisition repeated eighttimes for the reproducibility analysis.

Acquisition	kVp	mAs	Slice thickness (mm)	Pitch	Rotation time (s)	DFOV (mm)	ASIR (%)	Filter
Base	120	150	1.25	1.375	0.6	360	0	Standard
2	80	150	1.25	1.375	0.6	360	0	Standard
3	100	150	1.25	1.375	0.6	360	0	Standard
4	140	150	1.25	1.375	0.6	360	0	Standard
5	120	100	1.25	1.375	0.6	360	0	Standard
6	120	200	1.25	1.375	0.6	360	0	Standard
7	120	250	1.25	1.375	0.6	360	0	Standard
8	120	300	1.25	1.375	0.6	360	0	Standard
9	120	350	1.25	1.375	0.6	360	0	Standard
10	120	400	1.25	1.375	0.6	360	0	Standard
11	120	500	1.25	1.375	0.6	360	0	Standard
12	120	150	0.625	1.375	0.6	360	0	Standard
13	120	150	2.5	1.375	0.6	360	0	Standard
14	120	150	3.75	1.375	0.6	360	0	Standard
15	120	150	5	1.375	0.6	360	0	Standard
16	120	150	1.25	0.516	0.6	360	0	Standard
17	120	150	1.25	0.984	0.6	360	0	Standard
18	120	150	1.25	1.375	0.4	360	0	Standard
19	120	150	1.25	1.375	0.5	360	0	Standard
20	120	150	1.25	1.375	0.7	360	0	Standard
21	120	150	1.25	1.375	0.8	360	0	Standard
22	120	150	1.25	1.375	0.9	360	0	Standard
23	120	150	1.25	1.375	1	360	0	Standard
24	120	150	1.25	1.375	0.6	160	0	Standard
25	120	150	1.25	1.375	0.6	260	0	Standard
26	120	150	1.25	1.375	0.6	460	0	Standard
27	120	150	1.25	1.375	0.6	360	50	Standard
28	120	150	1.25	1.375	0.6	360	10	Standard
29	120	150	1.25	1.375	0.6	360	20	Standard
30	120	150	1.25	1.375	0.6	360	30	Standard
31	120	150	1.25	1.375	0.6	360	40	Standard
32	120	150	1.25	1.375	0.6	360	60	Standard
33	120	150	1.25	1.375	0.6	360	70	Standard
34	120	150	1.25	1.375	0.6	360	80	Standard
35	120	150	1.25	1.375	0.6	360	90	Standard
36	120	150	1.25	1.375	0.6	360	100	Standard
37	120	150	1.25	1.375	0.6	360	0	Soft
38	120	150	1.25	1.375	0.6	360	0	Detail
39	120	150	1.25	1.375	0.6	360	0	Bone
40	120	150	1.25	1.375	0.6	360	0	Lung

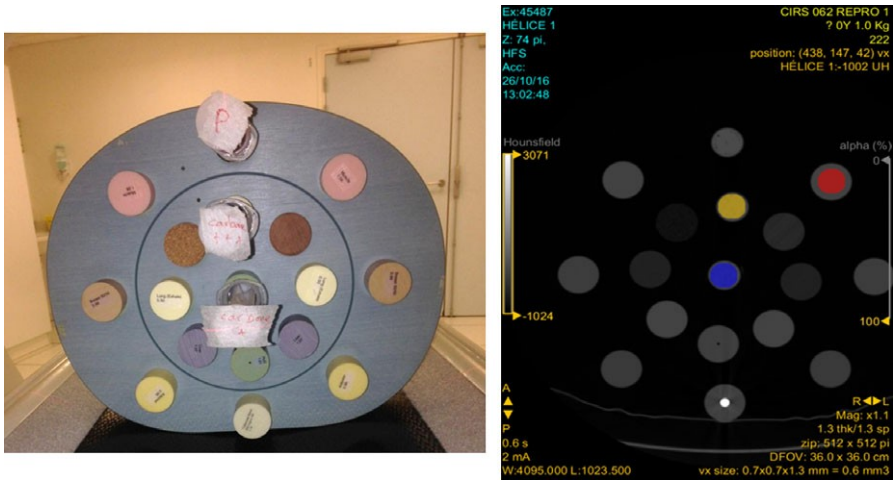

FIG. 2. Designed phantom with different inserts and resulting CT image of one slice of the phantom: 2D identical regions of interest were drawn on each of the three inserts: C1 for the homogeneous insert (upper right), C2 for the insert made of Ecoflex® + carbon1 (central circle), and C3 for the insert made of Ecoflex® + carbon2(above C2).

TABLE II. Descriptive values of each ROI in Hounsfield units for the eight identical CT acquisitions (1 to 8).

	CT acquisition							
	1	2	3	4	5	6	7	8
Name of ROI: C1								
minValue	25.0	21.0	20.0	14.0	23.0	22.0	20.0	24.0
maxValue	58.0	63.0	60.0	62.0	58.0	58.0	56.0	60.0
stdValue	6.0	6.3	6.2	6.3	6.4	5.9	5.7	5.9
meanValue	41.5	42.5	38.8	38.8	41.4	39.8	38.4	41.5
Name of ROI: C2								
minValue	−39.0	−32.0	−52.0	−52.0	−57.0	−45.0	−52.0	−62.0
maxValue	267.0	278.0	272.0	267.0	267.0	268.0	265.0	273.0
stdValue	41.9	43.0	42.9	42.8	42.6	42.5	42.3	43.2
meanValue	166.2	164.9	167.1	167.0	165.6	165.9	166.4	165.0
Name of ROI: C3								
minValue	−238.0	−237.0	−247.0	−238.0	−218.0	−246.0	−244.0	−246.0
maxValue	265.0	265.0	271.0	258.0	265.0	266.0	264.0	267.0
stdValue	65.5	64.6	66.8	66.4	63.9	65.7	66.1	66.1
meanValue	156.2	157.7	156.5	156.5	157.5	155.6	156.8	156.4

Finally, redundant indices were identified with a 2 x 2 correlation matrix (See Table S1 in Data S1) and a principal component analysis derived from the 39 supplementary acquisitions described in Table II. Based on the analysis of the 2 x 2 correlation matrix, skewness and kurtosis were independent of the six other indices but were intercorrelated for C2 and C3, and not for C1. The other six indices correlated with various degrees: four indices (contrast of the co- occurrence matrix, contrast of the Neighborhood Gray-Leveldifference matrix, entropy, and dissimilarity — were all associated with correlation coefficients >0.8, even for C1.

SZE was highly correlated with these four indices for C2 and C3, but to a lesser extent for C1. ZLNU was highly correlated with the five other indices for C1, but to a lesser extentfor C2 and C3.

Principal component analyses performed separately for each phantom pattern (C1, C2, C3) confirmed the 2 x 2 correlation analyses (fig. S1a–c: plots of the 1st and 2nd principal components, see Data S1). The patterns of the

three graphs were very similar. Six indices (dissimilarity, contrast,contrast.1, entropy, ZLNU, and SZE) contribute almost equally (around 15%) to the first principal component. A “size effect” was observed, that is, highly positively correlated variables are on the same side of the axis, with these six indices on the right side of the first axis. The second principalfactor is determined by skewness and kurtosis whose cumulative contribution to the second axis is above 85% for C1, C2and C3. The figures show the percentage of variance due to each component and the contribution of each index.

DISCUSSION

The use of texture analyses in medical imaging is anattractive option, potentially allowing the development of a novel form of disease “biomarkers”. Several teams worldwidehave developed software for texture analysis and have published in vivo data in widely varying fields, supporting the potential for major advances in our understanding of diseases.[12]

It is currently very difficult to draw a coherent position or recommendations from this abundant literature especially in order to select the proper texture indices that could be relevant in clinical practice.2In particular, the reproducibility of the calculation of textural features is currently a subject of debate. Our study indicates that many textural indices show a lack of reproducibility, even under the same experimental conditions. We can hypothesize that the variations we observed are due to the complexity of helical CT principles which involve detector properties, filtered back projection algorithm, and intrinsic noise. We also observed higher fluctuations within the homogenous area C1 than within C2 and C3. This could be explained by the use of an absolute discretization pixel intensity by steps of 10 HU in our experimental conditions which would have a greater influence on the smaller range of values of C1 (around 40 HU) compared to C2 and C3 (both around 300 HU).

TABLE III. Coefficient of variation (CV) of the values of each index calcu-lated from the eight identical CT acquisitions for C1, C2, and C3 areas.

LIFEx	CV	CV	CV	
Name of ROI	C1	C2	C3	
minValue	15.05	−18.65	−3.72	
meanValue	3.60	0.47	0.40	
stdValue	3.49	0.90	1.38	
maxValue	3.66	1.51	1.27	
SkewnessH	−375.13	−4.44	−1.80	
KurtosisH	5.11	3.37	2.28	
EntropyH	1.00	0.54	0.27	
EnergyH	3.17	3.20	1.45	
Homogeneity	1.12	0.88	0.80	
Energy	9.63	4.02	2.43	
Contrast (co-occurrence)		6.14	2.49	3.00
Correlation	8.50	0.80	0.38	
Entropy	3.50	0.35	0.34	
Dissimilarity	3.92	0.84	1.30	
SRE	1.85	0.42	0.44	
LRE	10.70	2.29	2.25	
LGRE	0.30	0.17	0.18	
HGRE	0.30	0.17	0.14	
SRLGE	2.01	0.38	0.41	
SRHGE	1.73	0.54	0.54	
LRLGE	10.82	2.33	2.19	
LRHGE	10.58	2.25	2.30	
GLNU	2.34	1.22	2.48	
RLNU	4.23	1.59	1.54	
RP	2.53	0.67	0.59	
Coarseness	4.46	2.21	1.43	
Contrast (NGLDM)	26.22	3.98	5.99	
Busyness	−744.65	−4039.39	−202.31	
SZE	8.37	1.68	1.28	
LZE	12.21	11.87	17.87	
LGZE	0.22	0.28	0.31	
HGZE	0.22	0.27	0.25	
SZLGE	8.29	1.50	1.35	
SZHGE	8.46	1.91	1.44	
LZLGE	12.19	11.77	17.51	
LZHGE	12.24	11.94	18.14	

GLNUz	11.10	3.73	2.65
ZLNU	14.30	5.08	3.59
ZP	6.72	2.56	1.41

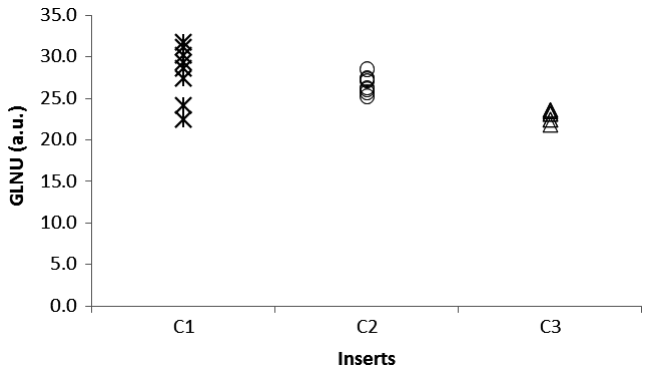

FIG. 3. Dot plot of the close distribution of values of the calculation of GLNU index from the Gray-Level Size Zone Matrix GLSZM of C1, C2, and C3.

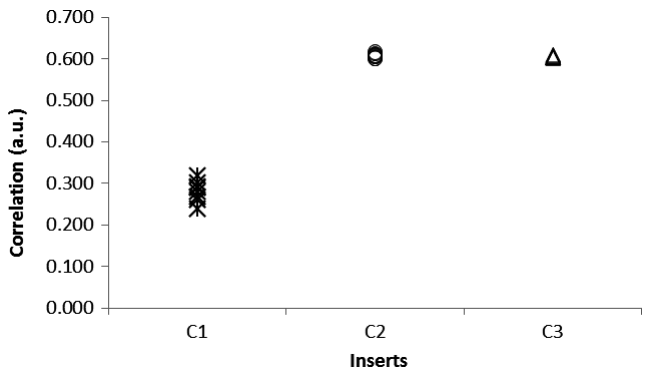

FIG. 4. Dot plot of the close distribution of values of the calculation of the correlation index between C2 and C3.

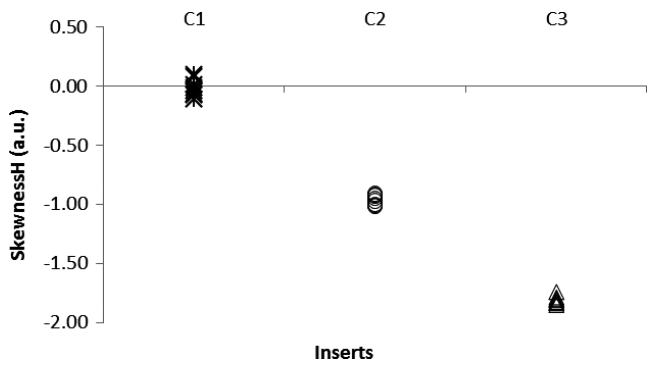

FIG. 5. Dot plot of the difference of the distribution of values of the calculation of skewness between the three ROI.

To our knowledge, this is the first study exploring the reproducibility of textural indices from CT images acquired eight times on a single device and a dedicated phantom. Some authors have already worried about this subject *in vivo*but the analysis of the literature is rendered difficult becausethe softwares and mathematical definition of each textural index used for textural indices often differ to some extent as no consensus currently exists.[12,13]
Fave et al.[14] studied the reproducibility of test–retest conebeam CT images of 10 patients explored with 68 textural features (histogram, co-occurrence matrix, run length matrix, NGLDM) computed with Ibex software. They excluded 23 features because they were not reproducible on the test-retestimages; unfortunately, with the exception of skewness derived from a histogram and contrast derived from the co-occurrence matrix, it is very difficult to compare their data with our study as we did not extract identical indices.
Hunter et al.[5] studied the reproducibility of test–retest of unenhanced CT images of 56 patients with lung tumors using three different devices (15, 16, and 25 patients for each device) explored with 328 textural features calculated by Ibex software (including histogram, co-occurrence matrixand gray-level run length matrix). They found a high concordance correlation coefficient (>0.90) between the two acquisitions for 61.0%–94.5% of the features, and a high concordance correlation coefficient for 138 image features when interpolating the reproducible features of each device. Of them, they found 23 non redundant features, including kurtosis, skewness derived from the histogram, and entropy derived from the co-occurrence matrix according to our results. Balagurunathan et al.[15] studied the reproducibilityof 219 3D textural indices derived from test–retest unenhanced CT scans of 32 patients with lung tumors and confirmed the robustness of contrast derived from the co-occurrence matrix.
Moreover, this is the first study to assess the usefulness ofindices for characterizing CT medical images on a dedicatedphantom. Our study shows that some textural indices are unable to discriminate homogeneous from heterogeneous features, likely rendering these indices unfeasible for use in clinical evaluation. Taking into account that our experimental conditions did not evaluate the influence of the size of the ROI, some indices are unable to discriminate two heterogeneous ROIs.
Our study has a number of limitations. We studied the indices extracted from the histogram and four matrices (co- occurrence, Gray-Level Run Length Matrix, Neighborhood Gray-Level difference matrix and Gray-Level Size Zone Matrix) from a single software (LIFEx), which is not as exhaustive as some published research,[4,14,15] however, our findings are in agreement with data published by other groups in terms of the fact that only a few textural indices arelikely to be sufficiently robust to be handled in clinical practice. In addition, we did not evaluate the influence of the sizeof the ROI. In the study of Fave et al.[14] the authors excludedthe following features: energy of the histogram, correlation ofthe co-occurrence matrix, LRHGE, SRLGE of the RLM matrix, busyness and contrast of the NGLDM matrix becausetheir values were found to be volume dependent.
We did not evaluate the influence of 3D vs 2D evaluation,but 2D images are more robust to variability than 3D imagesin patient studies.[16] In clinical routine and especially in oncology, tumor images are often ill-defined and subject to approximation in their segmentation, hence tumor segmentation is an important step to address before analysis of texturalfeatures is possible. Finally, we chose a discretization of 200gray levels, following results from other groups with LIFEx software in PET-CT imaging.[9,17]
In conclusion, our preliminary study highlighted a need for close attention when using texture analysis as a tool to characterize CT images because changes in quantitation maybe due to internal variability rather than due to actual physio-pathological effects. Some textural indices appear to be broadly influenced by noise, whereas 8 of 34 indices such asskewness, kurtosis derived from histogram, contrast, andentropy from the co-occurrence matrix, dissimilarity and contrast from the Neighborhood Gray-Level difference matrix NGLDM, and SZE and ZLNU from the Gray-Level Size Zone Matrix GLSZM appear to be sufficiently reliable and capable of discriminating close textures on CT images. Further research should be carried out to confirm and extend these results.

[a)]Author to whom correspondence should be addressed. Electronic mail: caroline.caramella@gustaverousy.fr

REFERENCES

1. Castellano G, Bonilha L, Li LM, et al. Texture analysis of medical images. *Clin Radiol*. 2004;59:1061–1069.
2. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*.

2016;278:563–577.
3. Limkin EJ, Sun R, Dercle L, et al. Promises and challenges for the implementation of computationnal medical imaging (radiomics) in oncology. *Ann Oncol*. 2017;28:1191–1206.
4. Fave X, Cook M, Frederick A, et al. Preliminary investigation into sources of uncertainty in quantitative imaging features. *Comput MedicalImaging Graph*. 2015;44:54–61.
5. Hunter L, Krafft S, Stingo F, et al. High quality machine-robust image features: identification in non-small cell lung cancer computed tomogra-phy images. *Med Phys*. 2013;40:121916.
6. Zhao B, Tan Y, Tsai WY, et al. Exploring variability in CT characteriza- tion of tumors: a preliminary phantom study. *Transl Oncol*. 2014;7:88–93.
7. Mackin D, Fave X, Zhang L, et al. Measuring computed tomography scanner variability of radiomics features. *Invest Radiol*. 2015;50:757– 765.
8. Orlhac F, Soussan M, Chouahnia K, et al. 18F-FDG PET-derived textu-ral indices reflect tissue-specific uptake pattern in non-small cell lung cancer. *PLoS ONE*. 2015;10:e0145063.
9. Nioche C, Orlhac F, Boughdad S, et al. A freeware for tumor hetero- geneity characterization in PET, SPECT, CT, MRI and US to accelerateadvances in radiomics. *J Nucl Med*. 2017;58:1316.
10. O'Connor JP, Aboagye EO, Adams JE. Imaging biomarker roadmap forcancer studies. *Nat Rev Clin Oncol*. 2017 Mar;14:169–186.

11. Abdi H, Williams LJ. Principal component analysis. *Wiley Interdiscip Rev Comput Stat*. 2010;2:433–459.
12. Bashir U, Siddique MM, Mclean E, et al. Imaging heterogeneity in lungcancer: techniques, applications, and challenges. *AJR Am J Roentgenol*.2016;207:534–543.
13. Buvat I, Orlhac F, Soussan M. Tumor texture analysis in PET: where do we stand? *J Nucl Med*. 2015;56:1642–1644.
14. Fave X, Mackin D, Yang J, et al. Can radiomic features be reproduciblymeasured from CBCT images for patients with non small cell lung can-cer? *Med Phy*. 2015;42:6784–6797.
15. Balagurunathan Y, Gu Y, Wang H, et al. Reproducibility and Prognosisof Quantitative Features Extracted from CT Images. *Transl Oncol*. 2014;7:72–87.
16. Zhao B, Tan Y, Bell DJ, et al. Exploring intra- and inter-reader variabil- ity in uni-dimensional, bi-dimensional, and volumetric measurements ofsolid tumors on CT scans reconstructed at different slice intervals. *Eur J Radiol*. 2013;82:959–968.
17. Orlhac F, Soussan M, Chouahnia K, et al. 18F-FDG PET-derived textu-ral indices reflect tissue-specific uptake pattern in non-small cell lung cancer. *PLoS ONE*. 2015;10:e0145063.

18. Haralick R, Shanmugam K, Dinstein I. Textural features for image clas-sification. *IEEE Trans Syst Man Cybern Syst*. 1973;6:610–621.
19. Galloway M. Texture analysis using gray level run lengths. *Comput Graph Image Process*. 1975;4:172–179.
20. Amadasun M, King R. Textural features corresponding to textural prop-erties. *IEEE Trans Syst Man Cybern Syst*. 1989;19:1264–1274.
21. Thibault G. *Texture indexes and gray level size zone matrix application to cell nuclei classification*. INSERM UMR 910, Medical Genetic and Functional Genomic, Medical.