

RESEARCH ARTICLE

A Study on the Basic Criteria for Selecting Heterogeneity Parameters of F18-FDG PET Images

Attila Forgacs^{1,2*}, Hermann Pall Jonsson², Magnus Dahlbom³, Freddie Daver⁴, Matthew D. DiFranco⁵, Gabor Opposits², Aron K. Krizsan², Ildiko Garai^{1,2}, Johannes Czernin³, Jozsef Varga², Lajos Tron², Laszlo Balkay²

1 Scanomed Nuclear Medicine Center, Debrecen, Debrecen, Hungary, **2** Department of Nuclear Medicine, Faculty of Medicine, University of Debrecen, Hungary, **3** Ahmanson Biological Imaging Center, Department of Molecular and Medical Pharmacology, David Geffen School of Medicine at University of California at Los Angeles, California, United States of America, **4** Alfred Mann Institute for Biomedical Engineering, University of Southern California, Los Angeles, California, United States of America, **5** Quantitative Imaging and Medical Physics at Medical University of Vienna, Vienna, Austria

* attilaforgacs@gmail.com



OPEN ACCESS

Citation: Forgacs A, Pall Jonsson H, Dahlbom M, Daver F, D. DiFranco M, Opposits G, et al. (2016) A Study on the Basic Criteria for Selecting Heterogeneity Parameters of F18-FDG PET Images. PLoS ONE 11(10): e0164113. doi:10.1371/journal.pone.0164113

Editor: Kewei Chen, Banner Alzheimer's Institute, UNITED STATES

Received: April 14, 2016

Accepted: September 20, 2016

Published: October 13, 2016

Copyright: © 2016 Forgacs et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The authors received no specific funding for this work.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

Textural analysis might give new insights into the quantitative characterization of metabolically active tumors. More than thirty textural parameters have been investigated in former F18-FDG studies already. The purpose of the paper is to declare basic requirements as a selection strategy to identify the most appropriate heterogeneity parameters to measure textural features. Our predefined requirements were: a reliable heterogeneity parameter has to be volume independent, reproducible, and suitable for expressing quantitatively the degree of heterogeneity. Based on this criteria, we compared various suggested measures of homogeneity. A homogeneous cylindrical phantom was measured on three different PET/CT scanners using the commonly used protocol. In addition, a custom-made inhomogeneous tumor insert placed into the NEMA image quality phantom was imaged with a set of acquisition times and several different reconstruction protocols. PET data of 65 patients with proven lung lesions were retrospectively analyzed as well. Four heterogeneity parameters out of 27 were found as the most attractive ones to characterize the textural properties of metabolically active tumors in FDG PET images. These four parameters included Entropy, Contrast, Correlation, and Coefficient of Variation. These parameters were independent of delineated tumor volume (bigger than 25–30 ml), provided reproducible values (relative standard deviation < 10%), and showed high sensitivity to changes in heterogeneity. Phantom measurements are a viable way to test the reliability of heterogeneity parameters that would be of interest to nuclear imaging clinicians.

Introduction

Medical imaging methods have an increasing role in cancer diagnostics and the assessment of therapy responses. Imaging methods providing lesion quantification information substantially help the evaluation of the efficiency of a chosen treatment. Lesion volume is the simplest parameter that can be determined when using any medical imaging methods. For a more complex characterization of lesions, quantitative measures can be routinely used in case of some imaging methods. Computed Tomography (CT), Magnetic Resonance (MR), Single Photon Emission Computed Tomography (SPECT), and Positron Emission Tomography (PET) are the most frequently used tomographic techniques. When considering these techniques overall, one of the advantages of PET imaging is that it is highly quantitative. The pixel values are directly related to the radiopharmaceutical uptake of the investigated organs. In case of the F¹⁸-FDG scans, the standardized uptake values (SUV_{max}, SUV_{mean}, SUV_{peak}) are the most frequently calculated quantitative parameters in oncological applications. A large number of articles have been published during the last decade to establish and reveal which SUV values and what cut off levels are the most appropriate for different tumor types in therapy monitoring and staging of cancer patients. Both the volume analysis and the SUV are confounded by the Partial Volume Effect (PVE). This effect is due to the limited spatial resolution and relatively high noise contributions of the PET systems. Therefore, any additional quantitative parameter would improve the prognostic and diagnostic capabilities of PET.

Tumor textural analysis is a new research field of growing interest and might give new insights for the quantitative characterization of tumors [1–10]. It has been determined that heterogeneity of the metabolically active tumor volume is mostly associated with necrosis, hypoxia, cellular proliferation, and angiogenesis. A number of these processes are closely related to cancer aggressiveness and prognosis [11–13]. Accordingly, local and regional texture analysis may provide additional information about the selected tumor tissue. In recent years, more than thirty different textural parameters have been used and investigated in F¹⁸-FDG studies [14]. However, their clinical utility and reliability are still uncertain. Brooks et al. state that any of the heterogeneity parameters should not be presumed as useful and reliable, and that one should not use them for clinical studies before complete mathematical and methodological standardization analyses have been performed [15–18]. They also demonstrate that the heterogeneity measures could be very sensitive to the volume in the case of small (<45 ml) tumors because of the limited spatial resolution of current PET scanners. Hatt et al. also found similar behavior for volumes above 10 ml, although the provided textural information increases extensively at larger volumes [19]. In addition, Galavis et al. and Jianhua et al. reported large variations for several heterogeneity parameters due to different acquisition and reconstruction parameters [20,21]. Recently, Orhac et al. [9] performed a comprehensive analysis including 31 textural indices with 3 different human cancer types and revealed that many heterogeneity parameters are somewhat redundant with one another. They also found that the heterogeneity values might depend on the tumor segmentation methods and the number of the predefined grey levels, as well. As a result, they proposed the use of only 4 robust and independent parameters for further investigation: Homogeneity, Entropy, Short-Run Emphasis (SRE), or Zone Length Non-Uniformity.

Phantom measurements could clarify the reliability of any textural characterization, excluding the overall bias caused by biological variability of human studies [14,22–26]. To the best of our knowledge, a limited number of papers deal with phantom-based analysis of textural features. A specially designed phantom was used for a reliability study to calculate different textural indices producing heterogenic activity distribution [27]. Nyflot et al. investigated the effects of stochastic image acquisition noise on the quantitative performance of the

Table 1. Different settings of reconstruction methods for phantom measurements on Siemens Biograph mCT scanner.

Type of Reconstruction	TOF	TrueX	Gauss filter [mm]		Pixel size [mm]	
			4	5	4	3,13
A	-	-	-	+	+	-
B	+	-	-	+	+	-
C	-	-	-	+	-	+
D	-	+	-	+	+	-
E	-	-	+	-	+	-
F	+	+	-	+	+	-

The “+” and “-” denote “yes” or “not” respectively to the application of the indicated reconstruction option.

doi:10.1371/journal.pone.0164113.t001

heterogeneity parameters with the help of the standard NEMA Image Quality phantom [28]. Orlhac and her coworkers presented work regarding how the resampling approach affects the ability of textural indices, utilizing uniform phantom measurement in addition to patient data [29]. In this study we propose phantom measurements and analysis to help choosing appropriate parameters to quantify tumor heterogeneity with well-defined conditions (reconstruction method, acquisition, minimal volume, segmentation). The following essential criteria were applied for a set of heterogeneity parameters: volume-independence, reproducibility, and quantitative ability to express the degree of heterogeneity. Once establishing this criteria, the selection of the most promising heterogeneity parameters (HePs) out of 27 parameters is analyzed and discussed.

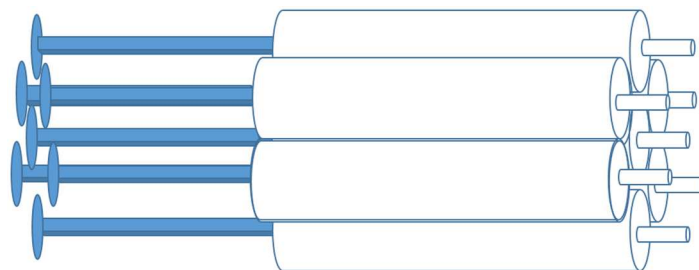
Methods and Materials

PET/CT Acquisitions and Reconstructions

Measurements were performed using Philips Gemini TF 64 (institute of author¹), GE Discovery ST 8 (VUE Point HD) (institute of author¹) and Siemens Biograph mCT (institute of author³) systems. The acquisitions and reconstructions were carried out using the default whole-body FDG-PET protocols for the Philips Gemini TF 64 and GE Discovery ST 8 scanner. On the Siemens Biograph system, a set of reconstructions were applied (see Table 1) using a variety of acquisition times in the range of 60 to 240 sec per bed position, including the factory default protocol.

Phantoms

A heterogeneous phantom insert (Revolver) was constructed using 7 pieces of 3ml syringes (inner diameter 8.66 mm) arranged in a revolver form (Fig 1). A 20 cm diameter standard

**Fig 1. Revolver heterogeneous phantom insert.**

doi:10.1371/journal.pone.0164113.g001

cylindrical phantom supplied by the manufacturer was also used to maintain homogeneous activity distribution, designated as the uniform phantom.

Patient Investigations

In addition to phantom measurements, PET data of 65 patients with confirmed lung lesions were retrospectively analyzed. Histological categorization of these tumors fell beyond our scope since focus was on the limitations of PET imaging. Patients fasted for at least 6 hours before the intravenous administration of 350–400 MBq of 18F-FDG. Blood glucose levels were always under 12 mmol/l. All patients were examined on the Philips Gemini TF scanner. Emission scans began 60 min after injection. Acquisition duration was 60–150 seconds per bed position, depending upon the patient weight. Delineation of lesion volumes was carried out by isocontouring with threshold at $SUV = 2.5 \text{ g/cm}^3$. Low dose CT scans were carried out without oral or intravenous administration of contrast agent. The study was approved by the local ethics committee (Regional And Institutional Ethics Committee, Clinical Center, University Of Debrecen). As the whole patient related data was a retrospectively analyzed, informed consent was not obtained.

Texture Parameters

All the images were evaluated using the Interview Fusion medical imaging software (Mediso Medical Imaging Systems Ltd., Budapest, Hungary). The computed heterogeneity parameters included Coefficient of Variation and the 26 derived matrix based heterogeneity parameters listed in Table 2. The proper equations are implemented based on Reference [30].

The voxel values of the segmented volume were resampled to yield a finite range of values allowing textural analysis using:

$$V(x) = \left[2^s \frac{I(x) - \min(I)}{\max(I) - \min(I) + 1} \right] \quad (1)$$

2^s represents the bit depth (in this study 64), $I(x)$ is the value of a given voxel of the original image [14].

Volume dependence

The homogeneous phantom was filled with 5 kBq/ml F-18 activity concentration and imaged according to the routine patient examination protocol on each scanner. Heterogeneity parameter values were calculated using concentric spherical Volume of Interests (VOIs) with volumes ranging from 0.5ml to 1000ml. The calculated parameters were then plotted versus VOI volumes.

Reproducibility

The reproducibility of the heterogeneity parameters was tested with the help of the heterogeneous phantom insert (Revolver insert) and NEMA IQ phantom. Revolver insert was placed into the NEMA IQ phantom (Fig 2A) filled with 5 kBq/ml F-18 activity concentration, while the syringes contained F-18 activity concentrations of 20 kBq/ml (blue syringes on Fig 2A), 40kBq/ml (green syringes on Fig 2A), and 80 kBq/ml (red syringes on Fig 2A), as displayed on Fig 2. This phantom setup was prepared and measured three times on the Siemens Biograph mCT scanner. Each of the three scan applied all the reconstructions listed in Table 1. in combinations of 60, 120, 180 and 240 sec/bed position.

Table 2. List of Indices Calculated from Texture Matrices, followed by the short name.

	Name of the heterogeneity parameter (HEP)
Co-occurrence matrix based indexes*	• Homogeneity- HOM;
	• Correlation- COR;
	• Entropy-ENT;
	• Contrast-CON
	• Intensity Variability-IV;
Volumetric Zone length statistics	• Zone Percentage- ZP;
	• Size-Zone Variability-SZV;
	• Short Zones Emphasis-SZE;
	• Long Zones Emphasis-LZE;
	• Grey-Level Non-Uniformity-GLNU _z
	• Low Grey-Level Zone Emphasis-LGLZE;
	• High Grey-Level Zone Emphasis- HGLZE;
	• Short Zone Low Grey Level Emphasis-SZLGLE;
	• Short Zone High Grey-Level Emphasis- SZHGLE;
	• Long Zone Low Grey Level Emphasis-LZLGLE;
	• Long Zone High Grey-Level Emphasis- LZHGLE;
Volumetric Run Length Statistics	• Run Percentage-RP;
	• Short Run Emphasis-SRE;
	• Long Run Emphasis- LRE;
	• Grey-Level Non-Uniformity- GLNU _R ;
	• Low Grey Level Run Emphasis-LGLRE;
	• High Grey Level Run Emphasis-HGLRE;
	• Short Run Low Grey-Level Emphasis- SRLGLE;
	• Short Run High Grey-Level Emphasis- SRHGLE;
	• Long Run Low Grey-Level Emphasis-LRLGLE;
	• Long Run High Grey-Level Emphasis-LRHGLE

*The co-occurrence type features were calculated for 26 different nearest neighbour connectivity and finally averaged over these directions.

doi:10.1371/journal.pone.0164113.t002

The high uptake volume imitating a lesion was defined by voxels with SUV>2.5 times the background within a boundary box. Mean value and standard deviation were calculated for each heterogeneity parameter from results of the 3 independent image sets. These calculations were carried out for each combination of acquisition times and reconstruction methods. The ratio of the standard deviation and mean (Coefficient of Variation) was used as the measure of error for the reproducibility of the selected parameter.

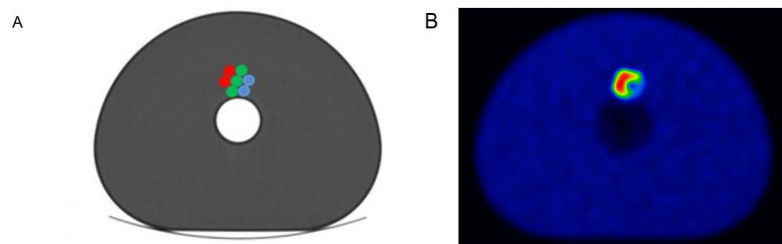


Fig 2. Revolver insert placed in the NEMA IQ Phantom. Illustrative schematic layout, (A) and a representative slice of the attenuation corrected PET image (B).

doi:10.1371/journal.pone.0164113.g002

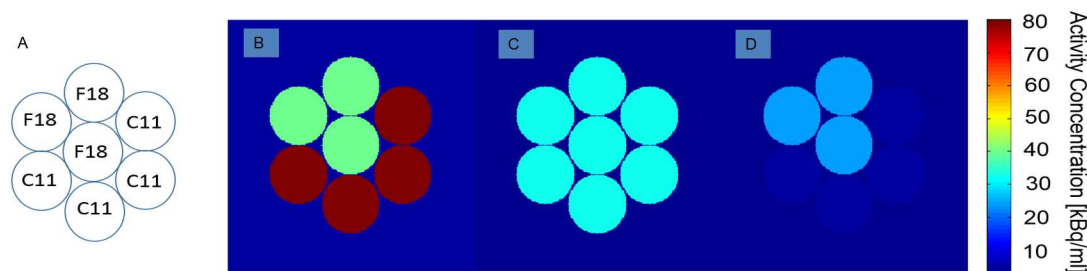


Fig 3. The geometry (A) and the activity distribution within the Revolver insert at (B) $t = 0$ min, (C) $t = 45$ min, and (D) $t = 80$ min.

doi:10.1371/journal.pone.0164113.g003

Parameter Sensitivity

Heterogeneity parameter values were determined for patterns with various degrees of heterogeneity. For this purpose, a part of Revolver insert was filled with C-11 solution (80 kBq/ml) and the remaining part of it with F-18 solution (40 kBq/ml). The Revolver insert was surrounded by F-18 water solution (5 kBq/ml) as homogeneous background.

PET scans were performed with 2min acquisitions each followed by a 2 min period of no acquisition. This was repeated over 20 cycles resulting in a total duration of 80 min on the Philips TF 64 scanner. Due to the difference in half-life between C-11 and F-18, the heterogeneity within the insert evolved over time as indicated in Fig 3. The dynamic data included an image at $t = 45$ min when the activity concentrations are equal for all syringes (Fig 3C). The VOI used for this set of measurements was defined by this image using a 2.5xBg threshold. Reconstruction was performed using the factory default protocol on Philips Gemini TF 64 scanner.

Results

Volume Dependence

Plots of the heterogeneity parameters versus the homogeneous volumes indicated that all of the texture parameters are volume dependent. Based on these results, we identified four different kinds of categories (Fig 4). The plots of volume dependence of the all investigated parameters can be found in S1 Fig.

In the first category, the volume dependence of the HeP values were limited to 25–30 ml volume range (Fig 4A). The second and third categories displayed volume dependence in the form of positive and negative slopes in log-log space, respectively (Fig 4B and 4C). The fourth category consisted of parameters which cannot be characterized by a simple function (Fig 4D). Table 3. summarizes the HePs and their categories.

The Fig 4 plots also demonstrate that the HeP data calculated from the human lung lesions and the homogenous phantom both fall in the similar ranges and follow the same behavior. This result emphasizes the relevance and importance of the simple homogenous phantom test to validate any promising HeP candidate.

Reproducibility

For parameters in the first category (first column in Table 3) the reproducibility errors (CV) were calculated and plotted (Fig 5) based on the heterogeneous phantom data from 3 independent scans. For further analysis we considered only those parameters where the CV values were less than 10% for almost all of the acquisition times and reconstruction settings. These parameters included Entropy, Homogeneity, Correlation, and Contrast.

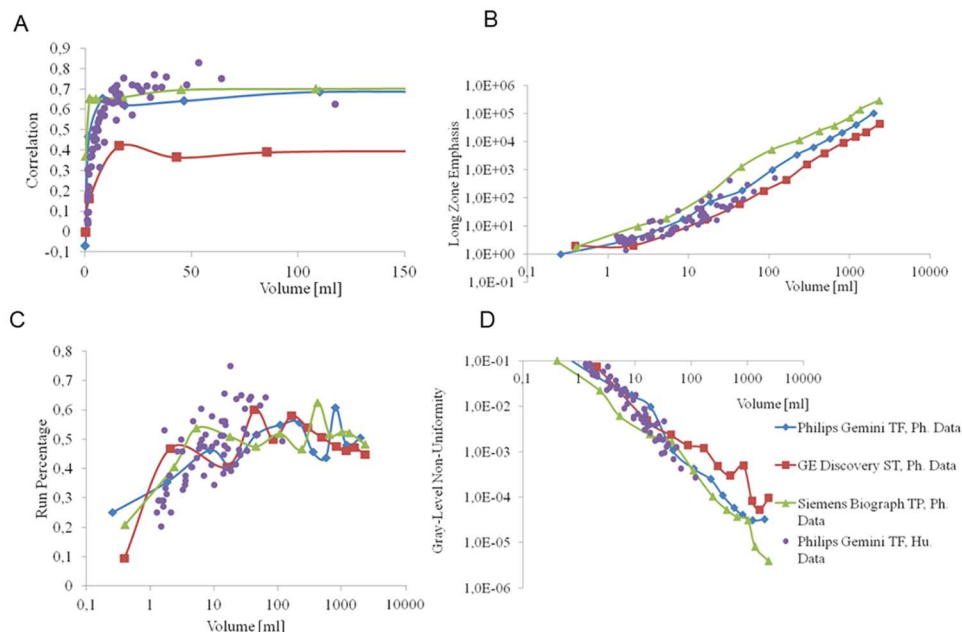


Fig 4. Representative volume dependence of four different HePs. Phantom data points ("Ph.Data") measured on three different scanners are differentiated by continuous color lines. Individual data points calculated from human lung lesions are displayed as the individual purple dots ("Hu.Data"). The volume dependence of the all investigated parameters can be found in the Supplemental Material.

doi:10.1371/journal.pone.0164113.g004

Coefficient of Variation

Calculation of the Coefficient of Variation parameter differs from HePs calculated using textural matrices. It does not require resampling and further matrix calculation; it is the ratio of deviation over mean. This difference in calculation method may account for the fact that the numerical values of Coefficient of Variation parameter evaluated from reconstructed human lung images differ in a magnitude from data derived from homogeneous phantom measurements (Fig 6A). Based on the results on Fig 6, the Coefficient of Variation parameter was included to the list of converging and reproducible parameters.

Parameter Sensitivity

At $t = 0$ min, the syringes filled with C-11 are far more active than those filled with F-18 as seen in Fig 7A. At $t = 45$ min, the activity concentrations of the syringes are relatively close resulting in a fairly even activity distribution throughout all syringes (Fig 7B). The boundary of the VOI applied for HeP calculations in this dynamic analysis is displayed on Fig 7B as well. At $t = 80$ min, the relatively high activity of the F-18 compared to the C-11 results in the image shown in Fig 7C. The heterogeneity parameters were calculated for the different time points, and are plotted on Fig 8.

Table 3. The Classification of 26 Textural Indices According to the Kind of Dependency of Parameter vs. Volume.

Converging (A)	Positive slope (B)	Random like (C)	Negative slope (D)
ENT, COR, HOM, CON, SZE, LGLZE, SZHGLE, HGLRE,	LZE, LZLGLE, LZHGLE, LRE	HGLZE, SZLGLE, LRLGLE, LRHGLE, RP	SZV, IV, GLNU _Z , GLNU _R , LGLRE, SRLGLE, SRHGLE, SRE, ZP,

doi:10.1371/journal.pone.0164113.t003

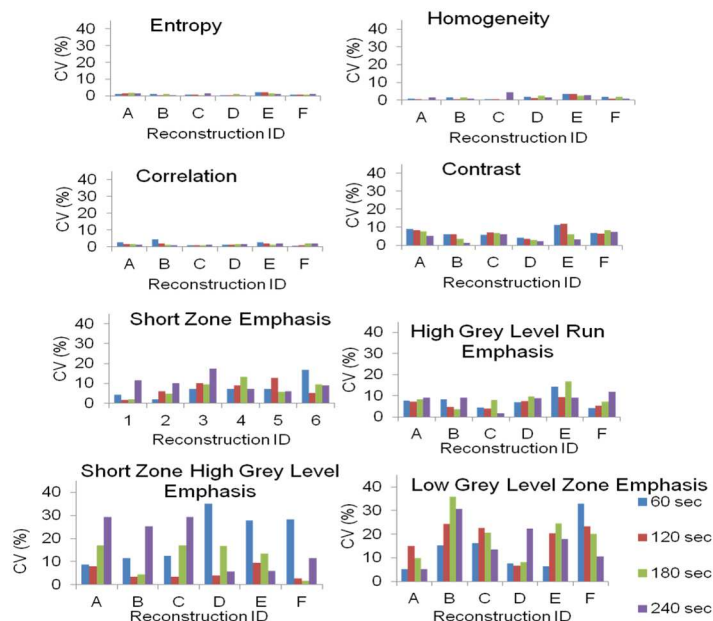


Fig 5. Reproducibility of the 8 remaining HePs (see Table 3, type A), as the function of reconstruction settings (see in Table 1) and acquisition time.

doi:10.1371/journal.pone.0164113.g005

Discussion

The main objective of this work was to quantitatively evaluate the reliability of a given set of heterogeneity parameters proposed for PET investigations. The heterogeneity of intratumoral radiopharmaceutical uptake has been suggested as a new quantitative measure and has been

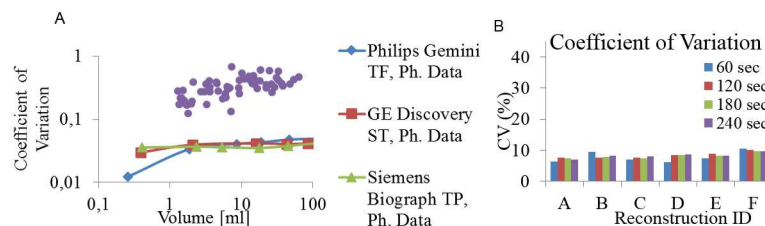


Fig 6. Volume dependence (A) and reproducibility (B) of the Coefficient of Variation parameter.

doi:10.1371/journal.pone.0164113.g006

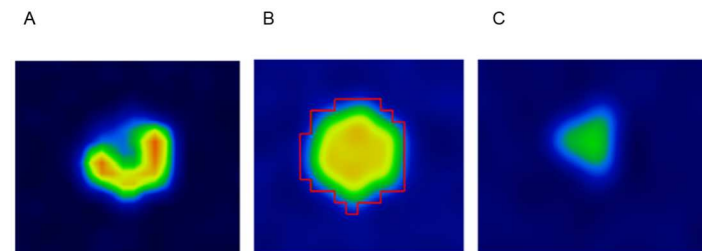


Fig 7. An axial slice of the image of Revolver insert phantom inserted into a homogeneous environment reconstructed from data of 2 min acquisition time beginning at $t = 0$ (A), $t = 45$ min (B), and $t = 80$ min (C). Boundary of the applied VOI is also displayed on panel b.

doi:10.1371/journal.pone.0164113.g007

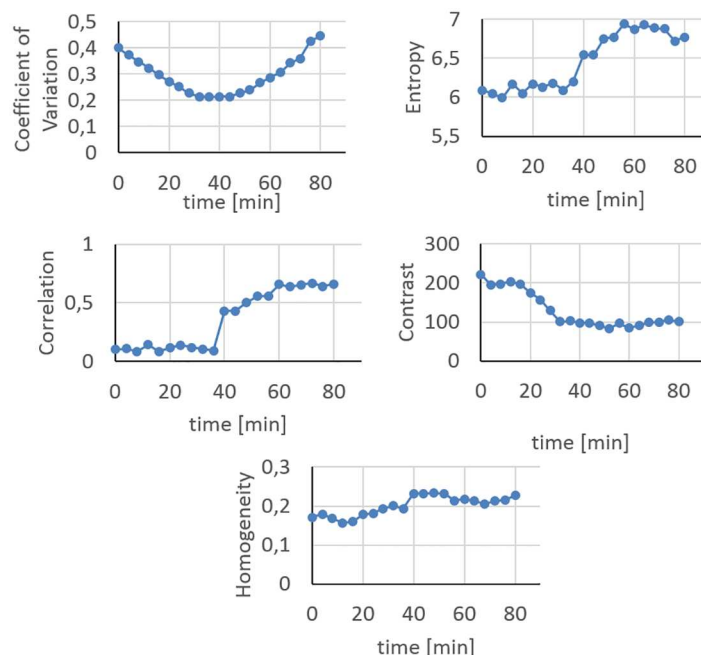


Fig 8. Time dependence of the heterogeneity parameters. Parameter values belonging to different time points are assigned to different textural patterns.

doi:10.1371/journal.pone.0164113.g008

intensively studied in a number of recent communications [2,4,5,8,31–36]. These parameters allow the quantification of intratumoral activity patterns that might reflect the metabolic and pathological state of the cancer. On the other hand, these parameters might help to predict the response of a chosen therapy for the patient. An increasing number of tumor heterogeneity studies have been published recently, particularly within the last two years. Review articles of this new field have been also released [1,3,5,37,38], supporting the growing interest of the heterogeneity measurements within the special field of cancer research. Moreover, any imaging based heterogeneity evaluation could lead to more personalized therapy.

Despite all of this great interest, a comprehensive verification of any of the heterogeneity parameters is still missing. Most articles present data about the unreliability of heterogeneity parameters. This unreliability is largely attributed to the lack of rigorous analysis of the effects of tumor volume, spatial resolution, reproducibility and image bit depth on these parameters [17,39]. The accuracy of the segmentation methods varies among different algorithms and also leads to increased reproducibility error [25,26]. To eliminate this variability, we chose the most robust segmentation method: the SUV threshold based segmentation. However, tumor delineation method of choice is a confounding issue, because different groups prefer different approaches. In addition, the defined volume that the texture indices are calculated within may influence the values [9]. More than 40 textural parameters have been proposed by different groups, but only a limited number of studies have clarified the differences among them, and tried to analyze the reliability and robustness of these parameters by applying sophisticated mathematical methods [9,15,34,40,41]. The comparative and follow up human studies have limited applicability for this purpose, because the biological and intersubjective variability may distort the basic characteristics of the selected heterogeneity parameter. In this work, we propose phantom-based methodologies to analyze the capability and reliability of textural parameters used with PET imaging. The tumor volume has to be large enough in order to reflect reliable heterogeneity. Heterogeneity parameters are the results of statistical calculations from

segmented voxels; these voxels affect each other due to the limited resolution of the system (i.e. point-spread function). When observing the behavior of a given HeP, the same type of volume dependence was observed in our phantom analysis (Fig 4) and confirmed by our retrospective lung image data. Based on these results we classified the 27 heterogeneity parameters into four clearly differentiated groups (Table 3). Due to their unpredictable volume dependence, the parameters categorized within the random-like group cannot offer useful characterization of texture. A larger number of HeP candidates show volume dependence similar to that displayed on Fig 4B and 4D. These parameters should be excluded from the further analysis as inadequate parameters. It can be strictly stated that well-performing heterogeneity parameters should not show dependence on the volume of the segmented region defined within the boundaries of a homogeneous phantom. Orhac et. al. [9] reported similar results based on patient data. They have found 10 or 19 parameters out of the investigated 31 as unsuitable if the volume correlation limit was set to $|r| > 0.8$ or $|r| > 0.6$, respectively. Based on our findings, 9 parameters out of the 27 showed volume independency above a minimal volume (25–30 ml; Table 3, Fig 4A). In addition, the human lung lesion data underscores the relevance of the phantom data, since the same behavior and range can be observed for both data sets in Fig 4. Brooks et al. examined the effect of small tumor volumes on the calculation of heterogeneity, using the combination of probability theory and clinical 18-FDG-PET data. They conclude that the heterogeneity parameters of tumor volumes below 45ml can be biased [15]. Reproducibility error was calculated for the following 9 well-performing parameters: Entropy, Contrast, Correlation, Homogeneity, Short Zone Emphasis, High Grey Level Run Emphasis, Short Zone High Grey Level Emphasis, Low Grey Level Zone Emphasis and Coefficient of Variation (Figs 5 and 6B). The reproducibility error of Entropy, Homogeneity and Correlation remained under 5% for each imaging protocol used, confirming their acceptable reliability. The error of Coefficient of Variation and Contrast parameters can be reduced below 10% by adequate selections of acquisition time and reconstruction method. The SZE, HGLRE, SZHGLE and the LGLZE were found to be less attractive parameters, and were excluded from further analysis because their error of reproducibility was higher than 10% in more than half of the imaging scenarios. PVE may quantitatively affect the voxel values, especially in the case of small patterns comparable to the spatial resolution. The PVE related behavior can be observed in Fig 4A, where in the range of small volumes (< 25 -30ml) the HeP was highly correlated with the volume. In addition, we also present PVE related data in Fig 5, where the reproducibility was investigated as the function of the reconstruction method and the acquisition time, since D and F reconstruction methods applied PVE correction (the so-called ‘TrueX’ option, see Table 1). However the TrueX option did not have great impact on neither the absolute value (S2 Fig) nor the reproducibility of the parameters.

We also constructed a dynamically changing texture by means of a dual isotope phantom. This phantom included separate syringes for F-18 and C-11 isotope and was used to investigate the sensitivity of the remaining HePs (Fig 7). As a result, an initially heterogeneous activity concentration pattern dominated by the C-11 is later transformed to an F-18 dominated heterogeneous pattern, across an equilibrium point. In this equilibrium point (at $t = 45$ min) the C-11 and F-18 activity distribution is equal, producing a fully homogeneous uptake pattern. Fig 8 demonstrates how the Entropy, Coefficient of Variation, Homogeneity, Correlation and Contrast differ in sensitivity and tendency.

The Coefficient of Variation parameter has two maxima according to the initial and the final heterogeneous uptake, and a minimum point at the equilibrium phase. Correlation has the highest sensitivity since the values are changing on a wide scale from 0.1 to 0.7. However, correlation, Entropy and Contrast moderately distinguish between the heterogeneous and homogeneous state. The Homogeneity parameter is a less sensitive parameter since it poorly

reflects the degree of heterogeneity. Therefore, it was excluded from the more reliable parameters. The remaining parameters to characterize the textural properties of metabolically active tumors in FDG PET scans which fulfill our criteria are the Entropy, Contrast, Correlation and Coefficient of Variation. These four parameters were selected from 27 candidates, however our selection strategy can be easily applied for any currently existing or newly introduced HeP, as well. Since our selection process was focused on physical validation, the investigation of the inherent prognostic value reflected by the validated parameters needs further human studies with multiple patient cohorts.

Although insightful, one of the primary limitations of this study is the lack of validation regarding the software used for the HeP calculations. In addition, several previously published studies used custom-written software without validation or comparison of the calculated HEP values with other software. Adequate phantom measurements could assist in this validation and comparison task for the whole imaging procedure including the acquisition, reconstruction, segmentation and HeP evaluation.

Conclusions

The purpose of the paper was to clearly define the criteria used to identify the most appropriate heterogeneity parameters to measure textural features. Based on this criteria, we compared various suggested measures of homogeneity. Our predefined criteria were: a reliable heterogeneity parameter has to be volume independent, reproducible and suitable to express quantitatively the degree of heterogeneity. Detailed phantom measurements and evaluation have been described to reveal the reliability of any existing or newly emerging parameters used to measure textural properties of delineated regions defined on reconstructed PET images. Only four out of the 27 selected textural parameters fully met our criteria. These parameters were the Coefficient of Variation, Contrast, Correlation and Entropy. By using appropriate reconstruction methods, these parameters can provide reproducible values within 10 percent error in tumor volumes > 25–30 ml. Our sensitivity study concluded that these four selected parameters have different levels of robustness to measure the degree of heterogeneity.

Supporting Information

S1 Fig. Volume dependence of all investigated HePs.

(PDF)

S2 Fig. The real value of the four finally suggested HePs as the function of reconstruction settings (see in [Table 1](#)) and acquisition time.

(TIF)

Acknowledgments

The authors would like to express special thanks to Mihaly Zentai, Viktoria Nagy and Dora Lili Pap for the thoughtful conversations and the helpful support during the period of this work.

Author Contributions

Conceptualization: AF LB.

Data curation: AF LB.

Formal analysis: AF LB.

Funding acquisition: AF LB JV.

Investigation: AF LB AK HPJ GO MD.

Methodology: AF LB.

Project administration: AF LB.

Resources: AF LB JC MD IG.

Software: AF LB.

Supervision: LB LT.

Validation: AF LB.

Visualization: AF LB.

Writing – original draft: AF LB FD MDD LT.

Writing – review & editing: AF LB FD MDD JV LT.

References

1. Alic L, Niessen WJ, Veenland JF. Quantification of Heterogeneity as a Biomarker in Tumor Imaging: A Systematic Review. PLoS One [Internet]. 2014; 9(10):e110300. Available: <http://dx.plos.org/10.1371/journal.pone.0110300> PMID: 25330171
2. Bundschuh R a, Dinges J, Neumann L, Seyfried M, Zsótér N, Papp L, et al. Textural Parameters of Tumor Heterogeneity in 18F-FDG PET/CT for Therapy Response Assessment and Prognosis in Patients with Locally Advanced Rectal Cancer. J Nucl Med [Internet]. 2014; 55:891–7. Available: <http://www.ncbi.nlm.nih.gov/pubmed/24752672> doi: 10.2967/jnumed.113.127340 PMID: 24752672
3. Buvat I, Orlhac F, Soussan M. Tumor Texture Analysis in PET: Where Do We Stand? J Nucl Med [Internet]. 2015 Nov 1; 56(11):1642–4. Available: <http://jnm.snmjournals.org/cgi/doi/10.2967/jnumed.115.163469> PMID: 26294296
4. Cheng N-M, Dean Fang Y-H, Tung-Chieh Chang J, Huang C-G, Tsan D-L, Ng S-H, et al. Textural Features of Pretreatment 18F-FDG PET/CT Images: Prognostic Significance in Patients with Advanced T-Stage Oropharyngeal Squamous Cell Carcinoma. J Nucl Med [Internet]. 2013; 54(10):1703–9. Available: <http://jnm.snmjournals.org/cgi/doi/10.2967/jnumed.112.119289> PMID: 24042030
5. Chicklore S, Goh V, Siddique M, Roy A, Marsden PK, Cook GJR. Quantifying tumour heterogeneity in 18F-FDG PET/CT imaging by texture analysis. Eur J Nucl Med Mol Imaging. 2013; 40(1):133–40. doi: 10.1007/s00259-012-2247-0 PMID: 23064544
6. Cook GJR, Yip C, Siddique M, Goh V, Chicklore S, Roy A, et al. Are Pretreatment 18F-FDG PET Tumor Textural Features in Non-Small Cell Lung Cancer Associated with Response and Survival After Chemoradiotherapy? J Nucl Med. 2013; 54(1):19–26. doi: 10.2967/jnumed.112.107375 PMID: 23204495
7. Dong X, Xing L, Wu P, Fu Z, Wan H, Li D, et al. Three-dimensional positron emission tomography image texture analysis of esophageal squamous cell carcinoma. Nucl Med Commun. 2013; 34(1):40–6. doi: 10.1097/MNM.0b013e32835ae50c PMID: 23111378
8. Nakajima EC, Laymon C, Oborski M, Hou W, Wang L, Grandis JR, et al. Quantifying Metabolic Heterogeneity in Head and Neck Tumors in Real Time: 2-DG Uptake Is Highest in Hypoxic Tumor Regions. PLoS One [Internet]. 2014; 9(8):e102452. Available: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4134191&tool=pmcentrez&rendertype=abstract> doi: 10.1371/journal.pone.0102452 PMID: 25127378
9. Orlhac F, Soussan M, Maisonnobe 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 [Internet]. 2014; 55(3):414–22. Available: <http://jnm.snmjournals.org/cgi/doi/10.2967/jnumed.113.129858> PMID: 24549286
10. Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Haibe-kains B, Grossmann P, et al. imaging using a quantitative radiomics approach. 2014;
11. Lu X, Kang Y. Hypoxia and Hypoxia-Inducible Factors: Master Regulators of Metastasis. Clin Cancer Res [Internet]. 2010 Dec 15; 16(24):5928–35. Available: <http://clincancerres.aacrjournals.org/cgi/doi/10.1158/1078-0432.CCR-10-1360> PMID: 20962028

12. Rajendran JG, Schwartz DL, O'Sullivan J, Peterson LM, Ng P, Scharnhorst J, et al. Tumor Hypoxia Imaging with [F-18] Fluoromisonidazole Positron Emission Tomography in Head and Neck Cancer. *Clin Cancer Res* [Internet]. 2006 Sep 15; 12(18):5435–41. Available: <http://clincancerres.aacrjournals.org/cgi/doi/10.1158/1078-0432.CCR-05-1773> PMID: 17000677
13. Kunkel M, Reichert TE, Benz P, Lehr H-A, Jeong J-H, Wieand S, et al. Overexpression of Glut-1 and increased glucose metabolism in tumors are associated with a poor prognosis in patients with oral squamous cell carcinoma. *Cancer*. 2003; 97:1015–24. doi: [10.1002/cncr.11159](https://doi.org/10.1002/cncr.11159) PMID: 12569601
14. Tixier F, Le Rest CC, Hatt M, Albarghach N, Pradier O, Metges J-P, et al. Intratumor Heterogeneity Characterized by Textural Features on Baseline 18F-FDG PET Images Predicts Response to Concomitant Radiochemotherapy in Esophageal Cancer. *J Nucl Med* [Internet]. 2011; 52(3):369–78. Available: <http://jnm.snmjournals.org/cgi/doi/10.2967/jnumed.110.082404> PMID: 21321270
15. Brooks FJ, Grigsby PW. The Effect of Small Tumor Volumes on Studies of Intratumoral Heterogeneity of Tracer Uptake. *J Nucl Med* [Internet]. 2014; 55(1):37–42. Available: <http://jnm.snmjournals.org/cgi/doi/10.2967/jnumed.112.116715> PMID: 24263086
16. Brooks FJ, Grigsby PW. FDG uptake heterogeneity in FIGO IIb cervical carcinoma does not predict pelvic lymph node involvement. *Radiat Oncol* [Internet]. 2013; 8(1):294. Available: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3904206&tool=pmcentrez&rendertype=abstract> doi: [10.1186/1748-717X-8-294](https://doi.org/10.1186/1748-717X-8-294) PMID: 24365202
17. Brooks FJ. On some misconceptions about tumor heterogeneity quantification. *Eur J Nucl Med Mol Imaging*. 2013; 40(8):1292–4. doi: [10.1007/s00259-013-2430-y](https://doi.org/10.1007/s00259-013-2430-y) PMID: 23632962
18. Brooks FJ, Grigsby PW. Current measures of metabolic heterogeneity within cervical cancer do not predict disease outcome. *Radiat Oncol* [Internet]. BioMed Central Ltd; 2011; 6(1):69. Available: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3130664&tool=pmcentrez&rendertype=abstract> doi: [10.1186/1748-717X-6-69](https://doi.org/10.1186/1748-717X-6-69) PMID: 21658258
19. Hatt M, Majdoub M, Valli M, Tixier F, Cheze C, Rest L, et al. PET Uptake Characterization Through Texture Analysis: Investigating the Complementary Nature of Heterogeneity and Functional Tumor Volume in a Multi-Cancer Site Patient Cohort. 2015; 56(1):38–45. doi: [10.2967/jnumed.114.144055](https://doi.org/10.2967/jnumed.114.144055) PMID: 25500829
20. Galavis PE, Hollensen C, Jallow N, Paliwal B, Jeraj R. Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters. *Acta Oncol* [Internet]. 2010; 49(7):1012–6. Available: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4091820&tool=pmcentrez&rendertype=abstract> doi: [10.3109/0284186X.2010.498437](https://doi.org/10.3109/0284186X.2010.498437) PMID: 20831489
21. Yan J, Chu-Shern JL, Loi HY, Khor LK, Sinha AK, Quek ST, et al. Impact of Image Reconstruction Settings on Texture Features in 18F-FDG PET. *J Nucl Med* [Internet]. 2015 Nov 1; 56(11):1667–73. Available: <http://jnm.snmjournals.org/cgi/doi/10.2967/jnumed.115.156927> PMID: 26229145
22. Weber WA, Ziegler SI, Thödtmann R, Hanauske AR, Schwaiger M. Reproducibility of metabolic measurements in malignant tumors using FDG PET. *J Nucl Med* [Internet]. 1999 Nov; 40(11):1771–7. Available: <http://www.ncbi.nlm.nih.gov/pubmed/10565769> PMID: 10565769
23. Nahmias C, Wahl LM. Reproducibility of standardized uptake value measurements determined by 18F-FDG PET in malignant tumors. *J Nucl Med*. 2008; 49(11):1804–8. doi: [10.2967/jnumed.108.054239](https://doi.org/10.2967/jnumed.108.054239) PMID: 18927325
24. Kinahan PE, Fletcher JW. Positron emission tomography-computed tomography standardized uptake values in clinical practice and assessing response to therapy. *Semin Ultrasound CT MR* [Internet]. 2010 Dec; 31(6):496–505. Available: <http://www.ncbi.nlm.nih.gov/pubmed/21147377> doi: [10.1053/j.sult.2010.10.001](https://doi.org/10.1053/j.sult.2010.10.001) PMID: 21147377
25. Frings V, de Langen AJ, Smit EF, van Velden FHP, Hoekstra OS, van Tinteren H, et al. Repeatability of Metabolically Active Volume Measurements with 18F-FDG and 18F-FLT PET in Non-Small Cell Lung Cancer. *J Nucl Med* [Internet]. 2010 Dec 1; 51(12):1870–7. Available: <http://jnm.snmjournals.org/cgi/doi/10.2967/jnumed.110.077255> PMID: 21078791
26. Hatt M, Cheze-Le Rest C, Aboagye EO, Kenny LM, Rosso L, Turkheimer FE, et al. Reproducibility of 18F-FDG and 3'-deoxy-3'-18F-fluorothymidine PET tumor volume measurements. *J Nucl Med* [Internet]. 2010; 51(9):1368–76. Available: <http://www.ncbi.nlm.nih.gov/pubmed/20720054> doi: [10.2967/jnumed.110.078501](https://doi.org/10.2967/jnumed.110.078501) PMID: 20720054
27. Forgacs A., Krizsan A. K., Garai I. et al. Reliability study of the calculated textural parameters for heterogenic activity distribution in PET investigation using special designed phantom. *Eur J Nucl Med Mol Imaging*. 2013; 40:S407–S407.
28. Nyflot MJ, Yang F, Byrd D, Bowen SR, Sandison GA, Kinahan PE. Quantitative radiomics: impact of stochastic effects on textural feature analysis implies the need for standards. 2015; 2(4):1–13. doi: [10.1117/1.JMI.2.4.041002](https://doi.org/10.1117/1.JMI.2.4.041002) PMID: 26251842

29. Orlhac F, Soussan M, Chouahnia K, Martinod E. 18F-FDG PET-Derived Textural Indices Reflect Tissue-Specific Uptake Pattern in Non-Small Cell Lung Cancer. 2015;1–16. doi: [10.1371/journal.pone.0145063](https://doi.org/10.1371/journal.pone.0145063) PMID: [26669541](https://pubmed.ncbi.nlm.nih.gov/26669541/)
30. SUPPLEMENTAL TABLE 1 Formulas for histogram and texture indices THE JOURNAL OF NUCLEAR MEDICINE • Vol. 55 • No. 3 • March 2014. 2014;55(3):2014.
31. Miwa K, Inubushi M, Wagatsuma K, Nagao M, Murata T, Koyama M, et al. FDG uptake heterogeneity evaluated by fractal analysis improves the differential diagnosis of pulmonary nodules. *Eur J Radiol* [Internet]. Elsevier Ireland Ltd; 2014; 83(4):715–9. Available: <http://www.ncbi.nlm.nih.gov/pubmed/24418285> doi: [10.1016/j.ejrad.2013.12.020](https://doi.org/10.1016/j.ejrad.2013.12.020) PMID: [24418285](https://pubmed.ncbi.nlm.nih.gov/24418285/)
32. Soussan M, Orlhac F, Boubaya M, Zelek L, Zioli M, Eder V, et al. Relationship between Tumor Heterogeneity Measured on FDG-PET/CT and Pathological Prognostic Factors in Invasive Breast Cancer. Singh PK, editor. *PLoS One* [Internet]. 2014 Apr 10; 9(4):e94017. Available: <http://dx.plos.org/10.1371/journal.pone.0094017> PMID: [24722644](https://pubmed.ncbi.nlm.nih.gov/24722644/)
33. Tixier F, Groves AM, Goh V, Hatt M, Ingrand P, Le Rest CC, et al. Correlation of Intra-Tumor 18F-FDG Uptake Heterogeneity Indices with Perfusion CT Derived Parameters in Colorectal Cancer. *PLoS One* [Internet]. 2014; 9(6):e99567. Available: <http://dx.plos.org/10.1371/journal.pone.0099567> PMID: [24926986](https://pubmed.ncbi.nlm.nih.gov/24926986/)
34. Tixier F, Hatt M, Valla C, Fleury V, Lamour C, Ezzouhri S, et al. Visual Versus Quantitative Assessment of Intratumor 18F-FDG PET Uptake Heterogeneity: Prognostic Value in Non-Small Cell Lung Cancer. *J Nucl Med* [Internet]. 2014; 55(8):1235–41. Available: <http://www.ncbi.nlm.nih.gov/pubmed/24904113> doi: [10.2967/jnumed.113.133389](https://doi.org/10.2967/jnumed.113.133389) PMID: [24904113](https://pubmed.ncbi.nlm.nih.gov/24904113/)
35. Yang F, Thomas MA, Dehdashti F, Grigsby PW. Temporal analysis of intratumoral metabolic heterogeneity characterized by textural features in cervical cancer. *Eur J Nucl Med Mol Imaging* [Internet]. 2013; 40(5):716–27. Available: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3625466&tool=pmcentrez&rendertype=abstract> doi: [10.1007/s00259-012-2332-4](https://doi.org/10.1007/s00259-012-2332-4) PMID: [23340594](https://pubmed.ncbi.nlm.nih.gov/23340594/)
36. Hatt M, Tixier F, Pierce L, Kinahan PE, Cheze C, Rest L, et al. Characterization of PET / CT images using texture analysis: the past, the present . . . any future? *Eur J Nucl Med Mol Imaging* [Internet]. European Journal of Nuclear Medicine and Molecular Imaging; 2016; Available: <http://dx.doi.org/10.1007/s00259-016-3427-0> PMID: [27271051](https://pubmed.ncbi.nlm.nih.gov/27271051/)
37. Asselin M-C, O'Connor JPB, Boellaard R, Thacker NA, Jackson A. Quantifying heterogeneity in human tumours using MRI and PET. *Eur J Cancer* [Internet]. Elsevier Ltd; 2012; 48(4):447–55. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0959804911010744> doi: [10.1016/j.ejca.2011.12.025](https://doi.org/10.1016/j.ejca.2011.12.025) PMID: [22265426](https://pubmed.ncbi.nlm.nih.gov/22265426/)
38. Davnall F, Yip CSP, Ljungqvist G, Selmi M, Ng F, Sanghera B, et al. Assessment of tumor heterogeneity: An emerging imaging tool for clinical practice? *Insights Imaging*. 2012; 3(6):573–89. doi: [10.1007/s13244-012-0196-6](https://doi.org/10.1007/s13244-012-0196-6) PMID: [23093486](https://pubmed.ncbi.nlm.nih.gov/23093486/)
39. Leijenaar RTH, Nalbantov G, Carvalho S, van Elmpt WJC, Troost EGC, Boellaard R, et al. The effect of SUV discretization in quantitative FDG-PET Radiomics: the need for standardized methodology in tumor texture analysis. *Sci Rep* [Internet]. 2015 Aug 5; 5:11075. Available: <http://www.nature.com/articles/srep11075> PMID: [26242464](https://pubmed.ncbi.nlm.nih.gov/26242464/)
40. Hatt M, Tixier F, Cheze Le Rest C, Pradier O, Visvikis D. Robustness of intratumour ¹⁸F-FDG PET uptake heterogeneity quantification for therapy response prediction in oesophageal carcinoma. *Eur J Nucl Med Mol Imaging* [Internet]. 2013; 40(11):1662–71. Available: <http://www.ncbi.nlm.nih.gov/pubmed/23857457> doi: [10.1007/s00259-013-2486-8](https://doi.org/10.1007/s00259-013-2486-8) PMID: [23857457](https://pubmed.ncbi.nlm.nih.gov/23857457/)
41. Brooks FJ, Grigsby PW. Quantification of heterogeneity observed in medical images. *BMC Med Imaging* [Internet]. 2013; 13:7. Available: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3682890&tool=pmcentrez&rendertype=abstract> doi: [10.1186/1471-2342-13-7](https://doi.org/10.1186/1471-2342-13-7) PMID: [23453000](https://pubmed.ncbi.nlm.nih.gov/23453000/)