Article:
Doumou, G, Siddique, M, Tsoumpas, C et al. (2 more authors) (2015) The precision of textural analysis in 18F-FDG-PET scans of oesophageal cancer. European Radiology, 25 (9). pp. 2805-2812. ISSN 0938-7994

https://doi.org/10.1007/s00330-015-3681-8

The precision of textural analysis in ^{18}F-FDG-PET scans of oesophageal cancer

Abstract

Objectives: Measuring tumour heterogeneity by textural analysis in ^{18}F-fluorodeoxyglucose positron emission tomography (^{18}F-FDG PET) provides predictive and prognostic information but technical aspects of image processing can influence parameter measurements. We therefore tested effects of image smoothing, segmentation and quantisation on the precision of heterogeneity measurements.

Methods: Sixty-four ^{18}F-FDG PET/CT scans of oesophageal cancer were processed using different Gaussian smoothing levels (2.0, 2.5, 3.0, 3.5, 4.0mm), maximum standardised uptake values (SUV_{max}) segmentation thresholds (45%, 50%, 55%, 60%) and quantisation (8, 16, 32, 64, 128 bin widths). Heterogeneity parameters included grey-level co-occurrence matrix (GLCM), grey-level run length matrix (GLRL), neighbourhood grey-tone difference matrix (NGTDM), grey-level size zone matrix (GLSZM) and fractal analysis methods. The Concordance Correlation Coefficient (CCC) for the 3 processing variables was calculated for each heterogeneity parameter.

Results: Most parameters showed poor agreement between different bin widths (median 0.08, range 0.004-0.99). Segmentation and smoothing showed smaller effects on precision (segmentation: median 0.82, range 0.33-0.97; smoothing: median 0.99, range 0.58-0.99).

Conclusions: Smoothing and segmentation have only a small effect on the precision of heterogeneity measurements in ^{18}F-FDG PET data. However, quantisation often has larger effects, highlighting a need for further evaluation and standardisation of parameters for multicentre studies.

Keywords: ^{18}F-FDG PET; Texture analysis; Heterogeneity; Precision; Oesophageal cancer

Key points:

1) Heterogeneity measurement precision in ^{18}F-FDG PET is influenced by image processing methods.

2) Quantisation shows large effects on precision of heterogeneity parameters in ^{18}F-FDG PET/CT.

3) Smoothing and segmentation show comparatively small effects on precision of heterogeneity parameters.

Introduction

^{18}F-fluorodeoxyglucose positron emission tomography (^{18}F-FDG-PET) significantly improves the accuracy of staging and therapy response assessment in a number of cancers [1,2]. There are early reports that textural analysis, an additional tool quantifying intratumoural heterogeneity of ^{18}FDG-PET tracer uptake, may improve prediction of response and prognosis and it is hypothesised that image heterogeneity may be related to underlying biology and reflect the behaviour of malignant tumours [3-7].

The measurement of tumour heterogeneity in ^{18}F-FDG PET images can be achieved by using statistical or model-based methods. Statistical-based textural analysis can be further categorised into first-, second- and higher-order statistical methods of increasing complexity, respectively [8-15]. The first-order statistical features are based on histograms of the original image. Second-order statistics describe the relationship between groups of two, usually neighbouring, voxels while high-order parameters, derived from 3D matrices, describe differences between each voxel and its neighbours, taking into consideration for each voxel, the neighbouring voxels in the two adjacent planes. For example, textural features from second-order statistics, grey-level co-occurrence matrices (GLCM), introduced by Haralick et al., describe the pixel distribution within a region and indicate the frequency of the appearance of various combinations of grey values [9]. The high-order neighbourhood grey-tone difference matrix (NGTDM) method computes the intensity differences between a voxel and its 26 neighbours [10]. Galloway first proposed the high-order grey-level run length (GLRL) matrix method that calculates the number of texels (run lengths) [11]. Texels are adjacent pixels with the same intensity. Chu et al. and Dasarathy and Holder added another two and four GLRL texture features, respectively [12,13]. High order grey-level size zone matrix (GLSZM) features were introduced by Thibault et al. as an extension of GLRL, giving information about the size and intensity of clusters of voxels or pixels in a region of interest [14]. Finally, model-based

fractal analysis (FA) methods describe the complexity of an object by identifying the property of self-similarity in the object itself [15,16].

Image segmentation is a factor that may depend on image noise and smoothing [17] and has potential effects on textural analysis [18]. Grey level quantisation (resampling to a number of bins) is an important process for the matrix construction. Since calculation of second or high order texture features from the large range of intensities within a PET image is computationally intensive, the data is typically binned, merging a large group of similar grey levels to a countable smaller number. The sampling ranges should be a finite number and thus bin widths as a power of 2 are chosen (8, 16, 32, 64, and 128). By including more levels in the bins, the extracted textural information will be more accurate and will result in a smoother image with reduced noise effects, but with consequent loss of information. Hence, the number of bin widths is a trade-off [19] and may influence textural feature measurements. To our knowledge, the sensitivity of textural features to different maximum standardised uptake value (SUV_{max}) segmentation thresholds, Gaussian smoothing levels and bin widths has not been evaluated together to date. The aim of our study was to evaluate the precision of textural feature measurements with respect to varying levels of these processing variables. In this study we did not aim to test the predictive or prognostic power of any of the texture variables which is the subject of a separate analysis.

Materials and methods

Dataset

Sixty-four patients with adenocarcinoma of the lower oesophagus underwent ^{18}F-FDG PET/CT scans for clinical staging purposes before surgery (n=64). Forty seven of the 64 patients were male and the mean age was 63.1 years. A waiver of institutional review board approval was obtained for this retrospective analysis.

Positron Emission Tomography Imaging

^{18}F-FDG PET/CT scans were all acquired as per standard institutional protocol on one of two scanners (Discovery VCT or DST, GE Healthcare, Waukesha, US) which are cross-calibrated to within 3% [20]. Patients were fasted for at least 6 hours prior to administration of 350-400 MBq ^{18}F-FDG. Scans were acquired 90 minutes after injection from the upper thigh to the base of skull for 4 minutes per bed position. Volumetric images were reconstructed using the ordered subset expectation maximisation (OSEM) algorithm (2 iterations, 20 subsets) with a slice thickness of 3.27mm and pixel size 4.7mm. Low dose CT was acquired for attenuation correction and anatomical localisation. The CT component of the scans was acquired at 120 kVp and 65 mAs without administration of oral or intravenous contrast agent.

Image analysis

In order to determine the effect of different Gaussian smoothing levels, percentage SUV_{max} segmentation thresholds and bin widths on the precision of texture features, different values of the associated variables were used, keeping the other parameters fixed (Table 1). Different Gaussian smoothing levels were added in the PET images by applying 2.0, 2.5, 3.0, 3.5 or 4.0 mm full width at half maximum (FWHM) Gaussian filters. Four different thresholds (45, 50, 55 or 60%) of percentage SUV_{max} were used to segment the primary oesophageal tumours by an experienced clinician. Finally, for the quantisation process, the following equation was used:

$$I_n = (N_g - 1)/(I_{max} - I_{min})\ (I - I_{min}) + 1 \qquad \text{(Equation 1)}$$

where N_g, is the value used for sampling the grey levels in different bin width ranges (8, 16, 32, 64, 128) and I, is the intensity.

Texture Analysis

After image processing, calculation of the textural features was performed using in-house software implemented under MATLAB (The MathWorks Inc.), constructing the matrices and calculating the 57

textural features from different matrices (GLCM, GLRL, GLSZM, NGTDM and FA). Table 2 lists the extracted features used in this study.

Statistical Analysis

Texture analysis measurements were statistically analysed by calculating the agreement between the different smoothing levels, bins or segmentation thresholds. We used the Concordance Correlation Coefficient (CCC) proposed by Lin (1989), as it has been shown to be an efficient calculation of agreement for multivariate and continuous data measured repeatedly by more than one method [21-23]. Moreover, it is suggested that CCC is not affected by outliers and scaling factors, in contrast to other agreement measurement methods [24].

One CCC was calculated for each of the pairs produced by combining the different parameters in each study (segmentation, smoothing, quantisation) (Table 1) and the CCC_{mean} of these pairs is presented to show the overall agreement between the altered parameters (Tables 3 and 4, and Appendix).

The scale below was used in order to classify the CCC scores [25]:

Value	Strength of Agreement
< 0.90	Poor
0.90 - 0.95	Moderate
0.95 - 0.99	Substantial
>0.99	Almost Perfect

As a result of the high segmentation percentages of 55% and 60% SUV_{max} thresholds, some tumour ROIs were divided into two new regions and therefore had to be excluded in the analysis process for the precision study.

Results

Smoothing

The mean CCC observed for most of the features with different levels of smoothing showed almost perfect agreement (37/57 textural features showed a CCC > 0.99) (Fig. 1). More specifically, GLCM, NGTDM and GLRL features showed the highest CCC scores with respect to different smoothing levels. The lowest scores were seen in GLSZM features which indicated the greatest effects from changes of smoothing levels (Tables 3 and 4). Only three out of thirteen GLSZM textural features presented substantial agreement (Short Zone Emphasis, Zone Percentage and Long Zone High Emphasis). Despite Fractal Dimension Mean (CCC = 0.85), fractal analysis techniques showed perfect agreement between different smoothing levels.

Segmentation

The mean CCC for the GLCM and GLRL textural features was slightly below 0.90, indicating only small effects from different segmentation thresholds on the measurement of these features (Fig. 2). However, most textural features derived from the high-order (NGTDM and GLSZM) and fractal analysis methods showed poor agreement between different bin widths. Substantial agreement was found in GLRL features Short Run High GL Intensity and High Run Emphasis, and a small number of textural features (7/57) showed moderate agreement within the range of 0.90-0.95.

Quantisation

Most of the features (51/57) showed low CCC scores (below 0.90) and 30 of them showed CCC ≥ 0.1 as a result of varying bin widths (Fig. 3). A minority of 6 out of 57 features showed CCC of higher than 0.90 (Tables 3 and 4). More specifically, perfect agreement was shown in Coarseness (NGTDM), substantial agreement in the GLCM features Correlation and Inverse Difference Moment Normalised, and moderate agreement in Lacunarity (FA), Short Run High GL Intensity (GLRL) and Inverse Difference Normalized (GLCM). Fractal analysis features were least affected by changes in bin width with mean CCC = 0.88 (Tables 3 and 4).

Discussion

To date there have been few data reported on the precision of ^{18}F-FDG PET texture features, i.e. the ability to obtain the same measurement from the scan data when changing parameters such as smoothing, segmentation and bin widths. This study evaluated the precision error of 57 texture features derived from ^{18}F-FDG PET images of oesophageal cancer with respect to different values of these three processing variables. The results show that changing smoothing levels has relatively small effects on the value of the majority of textural features, mostly demonstrating CCC values > 0.90. Similarly, changes in segmentation thresholds have small effects on most second-order and GLRL features but greater effects on high-order features. In contrast, changing the bin width produced poor agreements for most of the second- and high-order features, with the exception of fractal parameters. Overall, second-order features, as well as GLRL features, showed less sensitivity to changes in the three processing variables compared to the high-order features (Table 4). In particular, low CV% was observed for GLCM inverse difference moment normalised and inverse difference normalised, for GLRL Short Run High GL Intensity and High GL Run Emphasis, as well as for NGTDM Coarseness.. Moreover, entropy (GLCM) which has previously been reported as showing good test-retest reproducibility [26] and minimal sensitivity to various reconstruction parameters [27], showed CCC higher than 0.90 for smoothing and segmentation changes. Despite the large effect of segmentation on lacunarity and Fractal Dimension Standard Deviation, fractal analysis features otherwise were robust to smoothing, segmentation and bin width changes.

In a similar study of ^{18}F-FDG PET images in 3 cancer types, Orlhac et al. reported sensitivity of the majority (19/31) of first-, second- and higher-order features to segmentation methods [28] (40% of SUV_{max} vs Nestle method [29]). This was particularly true in some GLRL features compared to second-order GLCM features. There was also a marked effect from the resampling formula used and it was recommended that a bin width of at least 32 should be used to avoid introducing spurious relationships between texture features and SUV. In our study, NGTDM and GLSZM features were

particularly sensitive to varying the bin width. They also showed the lowest CCC when varying the SUV_{max} segmentation thresholds.

In contrast to the high sensitivity to segmentation and bin width changes seen with high-order statistical features in our study, a number of high-order regional features have shown good test-retest reproducibility similar to that found with SUV_{max} in another study of ^{18}F-FDG PET scans in patients with oesophageal carcinoma [30].

A further study examined the effects of different segmentation algorithms (fixed, adaptive and fuzzy locally adaptive Bayesian) and partial volume correction on textural features [18]. It was found that the calculated heterogeneity parameters were more sensitive to segmentation than partial volume correction. In general, second-order parameters, including entropy and homogeneity, were most robust.

In a study by Galavis et al., the raw ^{18}F-FDG-PET data of twenty patients diagnosed with different types of cancer were reconstructed with different acquisition modes and reconstruction parameters and some variability of textural features was noted [27]. In particular, the study evaluated the variability of 50 textural features between 2D and 3D acquisition modes and differing reconstruction algorithms and found that 40 of them showed large variations. The smallest variations were observed in energy, entropy (first-order), maximal correlation coefficient and low grey-level run emphasis with intermediate variation in entropy (second-order), sum entropy, high grey-level run emphasis and grey-level non-uniformity. These features that were included in our study were relatively robust to smoothing and segmentation changes.

A potential limitation of our study is that only fixed threshold methods of segmentation were used. Although our study showed that using thresholding as a segmentation method has little effect on the precision of most textural features, other techniques such as a fuzzy locally adaptive Bayesian (FLAB) algorithm [18] have been reported to lead to even smaller precision errors. In addition, only one formula was used to calculate the resampled values in our study and it has previously been noted that different resampling formulae can impact on texture feature calculations [28]. Similar to

previous studies, we only included ^{18}F-FDG PET scans of oesophageal carcinoma and it is possible that the effects of smoothing, segmentation and quantisation would be greater in other tumour types.

A variety of textural features have been described in medical imaging but it is known that there is often correlation between features [28] suggesting that the number of features used in future studies could be reduced. Whilst, in general, second-order, GLRL and fractal analysis parameters are the most robust with regards to the effects of smoothing and segmentation and some second- and high-order features have shown robustness in terms of test-retest reproducibility [26], acquisition mode and reconstruction methods [27], there is some variability in the strengths of individual parameters in the literature and selection of the number of bin widths would appear to be the dominant factor that requires optimisation and standardisation when considering the use of texture features in clinical practice or future studies.

Conclusion

There is growing interest in the measurement of intratumoural heterogeneity by textural analysis in PET and other imaging modalities as potential predictive and prognostic biomarkers. However, it is important that we understand the precision of these measurements and the effects of different processing and analytic methods before they become more widely used, particularly in the multicentre study setting. Whilst smoothing and segmentation methods have relatively small effects on most texture features, varying the bin width may have a significant effect on precision. Standardisation is key to successful clinical implementation of texture analysis.

Acknowledgements

The authors acknowledge support from the National Institute for Health Research Biomedical Research Centre of Guys & St Thomas' NHS Trust in partnership and King's College London and University College London Comprehensive Cancer Imaging Centre funded by Cancer Research UK and Engineering and Physical Sciences Research Council in association with the Medical Research Council and Department of Health (England).

References

1. Gambhir SS, Czernin J, Schwimmer J, Silverman DH, Coleman RE, Phelps ME (2001) A tabulated summary of the FDG PET literature. J Nucl Med 42:1S-93S

2. Krause BJ, Schwarzenböck S, Souvatzoglou M (2013) FDG PET and PET/CT. Recent Results Cancer Res 187:351-369

3. Tixier F, Le Rest CC, Hatt M, et al (2011) Intratumor heterogeneity characterized by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. J Nucl Med 52:369-378

4. Dong X, Xing L, Wu P, et al (2013) 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 34:40-46

5. Cook GJ, Yip C, Siddique M, Goh V, et al (2013) Are pretreatment 18F-FDG PET tumor textural features in non-small cell lung cancer associated with response and survival after chemoradiotherapy? J Nucl Med 54:19-26

6. Chicklore S, Goh V, Siddique M, Roy A, Marsden PK, Cook GJ (2013) Quantifying tumour heterogeneity in 18F-FDG PET/CT imaging by texture analysis. Eur J Nucl Med Mol Imaging 40:133-140

7. Davnall F, Yip CS, Ljungqvist G, et al (2012) Assessment of tumor heterogeneity: an emerging imaging tool for clinical practice? Insights Imaging 3:573-589

8. Castellano G, Bonilha L, Li LM, Cendes F (2004) Texture analysis of medical images. Clin Radiol 59:1061-1069

9. Haralick RM, Shanmugam K, Dinstein I (1973) Textural features for image classification. IEEE Trans Syst Man Cybern 3:610–621

10. Amadasun M, King R. Textural features corresponding to textural properties (1989) IEEE Trans Syst Man Cybern 19:1264–1274

11. Galloway MM. Texture analysis using gray level run lengths (1975) Comput Graph Image Process 4:172-179

12. Chu A, Sehgal C. M, Greenleaf JF (1990) Use of gray value distribution of run lengths for texture analysis. Pattern Recognit Lett 11:415–420

13. Dasarathy BR, Holder EB (1991) Image characterizations based on joint gray-level run-length distributions. Pattern Recognit Lett 12:497–502

14. Thibault G, Angulo J, Meyer F (2014) Advanced statistical matrices for texture characterization: application to cell classification. IEEE Trans Biomed Eng 61:630-637

15. Lopes R, Betrouni N. Fractal and multifractal analysis: a review (2009) Med Image Anal 13:634-649

16. Goh V, Sanghera B, Wellsted DM, Sundin J, Halligan S (2009) Assessment of the spatial pattern of colorectal tumour perfusion estimated at perfusion CT using two-dimensional fractal analysis. Eur Radiol 19:1358-1365

17. Tamal M (2012) Threshold based segmentation in Positron emission tomography for radiotherapy planning and treatment assessment. Current Mol Imaging 1:63-68

18. Hatt M, Cheze le Rest C, Turzo A, Roux C, Visvikis D (2009) A fuzzy locally adaptive Bayesian segmentation approach for volume determination in PET. IEEE Trans Med Imaging 28: 881-893

19. Soh LK, Tsatsoulis C (1999) Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. IEEE Trans Geosci Remote Sens 37:780-795

20. Schleyer PJ, Baker S, Barrington SF, et al (2008) Establishment of acquisition and reconstruction parameters for a GE Discovery VCT PET-CT scanner. Eur J Nucl Med Mol Imaging 35:S340-341

21. Lin LIK (1989) A concordance correlation coefficient to evaluate reproducibility. Biometrics 45:255–268

22. Strother SC, Lange N, Anderson JR, Schaper KA, Rehm K, Hansen LK and Rottenberg DA (1997) Activation pattern reproducibility: measuring the effects of group size and data analysis models. Hum. Brain Map. 5:312–316

23. Lange N, Strother SC, Anderson JR, Nielsen F, Holmes A, Kolenda T, Savoy R and Hansen L (1999) Plurality and resemblance in fMRI data analysis. NeuroImage, 10:282-303

24. Vrenken H, Vos EK, van der Flier WM, Sluimer IC, Cover KS, Knol DL, et al. (2013) Validation of the automated method VIENA: An accurate, precise, and robust measure of ventricular enlargement. Hum. Brain Map. 35:1101-1110

25. McBride GB (2005) A proposal for strength-of-agreement criteria for Lin's Concordance Correlation Coefficient. NIWA Client Report: HAM2005-062

26. Tixier F, Hatt M, Le Rest CC, Le Pogam A, Corcos L, Visvikis D (2012) Reproducibility of tumor uptake heterogeneity characterization through textural feature analysis in 18F-FDG PET. J Nucl Med 53:693-700

27. Galavis PE, Hollensen C, Jallow N, Paliwal B, Jeraj R (2010) Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters. Acta Oncologica 49:1012-1016

28. Orlhac F, Soussan M, Maisonobe JA, Garcia CA, Vanderlinden B, Buvat I (2014) 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 55:414-422

29. Nestle U, Kremp S, Schaefer-Schuler A, et al (2005) Comparison of different methods for delineation of 18F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non–small cell lung cancer. J Nucl Med 46:1342–1348

30. Hatt M, Tixier F, Cheze Le Rest C, Pradier O, Visvikis D (2013) Robustness of intratumour ^{18}F-FDG PET uptake heterogeneity quantification for therapy response prediction in oesophageal carcinoma. Eur J Nucl Med Mol Imaging 40:1662-1671

Figure legends

Fig. 1 Bar chart illustrating the CCC observed for all 57 textural features with different levels of smoothing. The features are presented with numbers according to table 2.

Fig. 2 Bar chart illustrating the CCC observed for all 57 textural features with different segmentation thresholds of percentage SUV_{max}. The features are presented with numbers according to table 2.

Fig. 3 Bar Chart illustrating the CCC observed for all 57 textural features with different bin widths. The features are presented with numbers according to table 2.

Tables

TABLE 1. Combinations of the parameters used in this study.

Condition	SUV_{max} (%)	Bin Width	Smoothing (mm)
Segmentation	**45**	16	2.5
	50	16	2.5
	55	16	2.5
	60	16	2.5
Smoothing	45	16	**2.0**
	45	16	**2.5**
	45	16	**3.0**
	45	16	**3.5**
	45	16	**4.0**
Bin Widths	45	**8**	2.5
	45	**16**	2.5
	45	**32**	2.5
	45	**64**	2.5
	45	**128**	2.5

SUV_{max} – maximum standardised uptake value

TABLE 2. Analytical table with texture parameters derived after statistical or model based analysis.

Method	**Order**	**Type**
Statistical based analysis	Second-order statistics	GLCM 1. Angular Second Moment 2. Autocorrelation 3. Cluster Prominence 4. Cluster shade 5. Contrast 6. Correlation 7. Difference Entropy 8. Difference Variance 9. Dissimilarity 10. Energy 11. Entropy 12. Homogeneity 13. Information Measure Correlation 1 14. Information Measure Correlation 2 15. Difference Moment 16. Inverse Difference Moment Normalised 17. Inverse Difference Normalised 18. Maximum Probability 19. Sum Average 20. Sum Entropy 21. Sum Variance 22. Sum Squares Variance
Statistical based analysis	Higher-order statistics	GLRL 23. Short Run Emphasis 24. Long Run Emphasis 25. Grey Level Non-uniformity 26. Run Length Non-uniformity 27. Run Percentage 28. Low Grey Level Run Emphasis 29. High Grey Level Run Emphasis 30. Short Run Low Grey Level Intensity 31. Short Run High Grey Level Intensity 32. Long Run High Grey Level Intensity 33. Intensity Variability 34. Run Length Variability 35. Long Run Low Grey Level Intensity

		NGDTM 36. Coarseness 37. Contrast 38. Busyness 39. Complexity 40. Texture Strength GLSZM 41. Short Zone Emphasis 42. Long Zone Emphasis 43. Intensity Non-uniformity 44. Zone Length Non-uniformity 45. Zone Percentage 46. Low Intensity Zone Emphasis 47. High Intensity Zone Emphasis 48. Short Zone Low Emphasis 49. Short Zone High Emphasis 50. Long Zone Low Emphasis 51. Long Zone High Emphasis 52. Intensity Variability 53. Size zone Variability
Model based analysis	Fractal Analysis	FA 54. Fractal Dimension Mean 55. Fractal Dimension Standard Deviation 56. Lacunarity 57. Hurst Exponent

GLCM – grey-level co-occurrence matrix, GLRL – grey-level run length, NGTDM - neighbourhood grey-tone difference matrix, GLSZM – grey-level size zone matrix, FA - fractal analys

TABLE 3. Concordance Correlation Coefficient for each texture feature depending on changes in smoothing, segmentation threshold and bin width.

Textural Feature	Smoothing CCC	Segmentation CCC	Bin width CCC
1. Angular Second Moment	0.99	0.93	0.05
2. Autocorrelation	0.99	0.87	0.01
3. Cluster Prominence	0.97	0.89	0.004
4. Cluster shade	0.99	0.79	0.02
5. Contrast	0.99	0.83	0.04
6. Correlation	0.99	0.71	0.98
7. Difference Entropy	0.99	0.78	0.04
8. Difference Variance	0.99	0.83	0.04
9. Dissimilarity	0.99	0.84	0.05
10. Energy	0.99	0.93	0.05
11. Entropy	0.99	0.92	0.05
12. Homogeneity	0.99	0.86	0.07
13. Information Measure Correlation 1	0.99	0.57	0.28
14. Information Measure Correlation 2	0.99	0.81	0.16
15. Difference Moment	0.99	0.86	0.07
16. Inverse Difference Moment Normalised	0.99	0.83	0.96
17. Inverse Difference Normalised	0.99	0.84	0.93
18. Maximum Probability	0.98	0.76	0.06
19. Sum Average	0.99	0.85	0.02
20. Sum Entropy	0.98	0.88	0.02
21. Sum Variance	0.99	0.87	0.01
22. Sum Squares Variance	0.99	0.86	0.01
23. Short Run Emphasis	0.95	0.81	0.16
24. Long Run Emphasis	0.99	0.89	0.01
25. Grey Level Non-uniformity	0.99	0.86	0.61
26. Run Length Non-uniformity	0.99	0.91	0.56
27. Run Percentage	0.99	0.84	0.26
28. Low GL Run Emphasis	0.99	0.93	0.19
29. High GL Run Emphasis	0.99	0.97	0.83
30. Short Run Low GL Intensity	0.96	0.79	0.16
31. Short Run High GL Intensity	0.99	0.97	0.91
32. Long Run Low GL Intensity	0.99	0.90	0.01
33. Long Run High GL Intensity	0.99	0.77	0.11
34. Intensity Variability	0.99	0.83	0.89
35. Run Length Variability	0.99	0.85	0.48

36. Coarseness	0.99	0.82	0.99
37. Contrast	0.99	0.80	0.31
38. Busyness	0.99	0.59	0.11
39. Complexity	0.97	0.33	0.01
40. Texture Strength	0.99	0.61	0.08
41. Short Zone Emphasis	0.95	0.46	0.02
42. Long Zone Emphasis	0.91	0.69	0.01
43. Intensity Non-uniformity	0.90	0.84	0.09
44. Zone Length Non-uniformity	0.92	0.71	0.61
45. Zone Percentage	0.97	0.92	0.38
46. Low Intensity Zone Emphasis	0.58	0.42	0.05
47. High Intensity Zone Emphasis	0.92	0.54	0.03
48. Short Zone Low Emphasis	0.82	0.47	0.03
49. Short Zone High Emphasis	0.63	0.54	0.01
50. Long Zone Low Emphasis	0.85	0.63	0.01
51. Long Zone High Emphasis	0.95	0.54	0.22
52. Intensity Variability	0.89	0.81	0.08
53. Size zone Variability	0.91	0.70	0.38
54. Fractal Dimension Mean	0.85	0.81	0.85
55. Fractal Dimension standard deviation	0.99	0.65	0.89
56. Lacunarity	0.99	0.42	0.93
57. Hurst Exponent	0.99	0.81	0.85

CCC – Concordance Correlation Coefficient, GL – grey-level

TABLE 4**.** Mean, median and range of the CCC for each group of texture features according to changes in smoothing, segmentation and bin widths

Type of texture feature	Smoothing Mean CCC	Median CCC	Range CCC	Segmentation Mean CCC	Median CCC	Range CCC	Bin width Mean CCC	Median CCC	Range CCC
GLCM	0.99	0.99	0.97-0.99	0.83	0.85	0.57-0.93	0.18	0.05	0.004-0.98
GLRL	0.98	0.99	0.95-0.99	0.87	0.86	0.77-0.97	0.40	0.26	0.01-0.91
NGTDM	0.99	0.99	0.97-0.99	0.63	0.61	0.33-0.82	0.30	0.11	0.01-0.99
GLSZM	0.86	0.91	0.58-0.97	0.64	0.63	0.42-0.92	0.15	0.05	0.01-0.61
FA	0.96	0.99	0.85-0.99	0.67	0.73	0.42 -0.81	0.88	0.87	0.85-0.93

CCC – Concordance Correlation Coefficient, GLCM – grey-level co-occurrence matrix, GLRL – grey-level run length, NGTDM - neighbourhood grey-tone difference matrix, GLSZM – grey-level size zone matrix, FA - fractal analysis.