

The Statistical Influence of Imaging Time and Segmentation Volume on PET Radiomic Features: A Preclinical Study

Emad M. Alsayed, Rhodri Smith, Christopher Marshall, Stephen Paisey, Emiliano Spezi

Abstract— Medical imaging plays an essential role in the diagnosis and treatment of many types of cancer. Currently, medical images are assessed visually by radiologists and clinicians. However, the full utility of information contained within medical images has yet to be fully explored. One avenue for this exploration is the utilization of “radiomic features” through the application of texture analysis. The numerous radiomic features proposed may vary with confounding variables such as the time post injection of image acquisition and the accuracy of the delineation of the prescribed segmentation volume. To this avail, we propose using the determinant of the correlation matrix to analyze radiomic features robustness to confounding variables. For this purpose, dynamic pre-clinical positron emission tomography (PET) images of 8 mice with mammary carcinoma xenografts (4T1) were binned into 5 minutes intervals from 50 to 70 minutes post injection. The effect of variation in segmentation was also explored by incrementally increasing segmentation volume. From each image set, we extracted 78 Radiomic features for analysis. Analysis. The statistical association measured by the determinant of the correlation matrix when considering contour size was 0.02378; for acquisition time this value was 0.13296. From this analysis we conclude that both temporal variation and segmentation effect the measurement of temporal features and that texture features are less robust to varying acquisition time than to varying segmentation volume.

Index Terms — PET, Preclinic, Radiomic, Cancer, Robustness

I. INTRODUCTION

MEDICAL imaging plays a vital role in the diagnosis and treatment for many types of cancer. Positron Emission Tomography (PET) imaging contributes significantly in the assessment and management of cancer [1]. Currently, PET images and other medical images are interpreted visually by radiologists and clinicians. However, medical images contain more information than can be assessed visually [2]. As a result, there is an increasing interest in artificial intelligence (AI) for automated image analysis [3], [4] and decision making [5].

Within this paradigm radiomics may be utilized to build predictive models to assess treatment outcome. It is now accepted that further data extraction has the potential to enhance the prognostic and diagnostic power of the radiologist or oncologist [6], [7]. Radiomics is defined as the field involved in extracting high dimensional quantitative features from medical images. However, radiomic features however may vary with different factors and conditions [8]. Previous research has highlighted numerous confounding factors that affect texture features such as image reconstruction type [7], respiratory motion [8] and variability in the delineation of the segmented volume [11], [12]. For increased confidence in the utilization of texture features as imaging biomarkers it is necessary to understand the extent these confounding factors have on Radiomic analysis which will inform attempts to standardize texture analysis. To the authors knowledge no published work has investigated and compared the impact of two variables (segmentation volume and acquisition time) on the stability or robustness of PET radiomic feature.

In this work, we evaluate using a novel approach the impact of segmentation volume and acquisition time on 18F-FDG PET radiomic features. To fully assess the aforementioned confounding variables influence, the determinant of the correlation matrices for each feature whilst varying segmentation contour sizes and image acquisition times was calculated. The determinant of the correlation matrix is related to the volume of the space occupied by the swarm of the standardized data points (mean = 0, standard deviation =1). When the measures are uncorrelated, this space is a sphere with a volume of 1. When the measures are correlated, the space occupied becomes an ellipsoid whose volume is less than 1. Thus, as the volume approaches 0 the variables are more correlated. If the determinant of the correlation matrix is equal to one this indicates the columns of the correlation matrix are orthonormal. If the determinant of the correlation matrix is equal to zero, this indicates the existence of exact linear dependence amongst the variables with the possibility for predictive analysis using a linear model.

Manuscript received December 30, 2019. E. M. Alsayed is supported by king Abdulaziz University, Jeddah, Saudi Arabia (grant number # KAU1938).

E. M. Alsayed is PhD student at School of Engineering, Cardiff University, Cardiff, CF24 3AA, UK (e-mail: alsyede@cardiff.ac.uk)

R. Smith is clinical scientist with Wales Research and Diagnostic PET Imaging Centre, Cardiff, CF14 4XN, UK (e-mail: Smithr50@cardiff.ac.uk)

C. Marshall is Director of Wales Research and Diagnostic PET Imaging Centre, Cardiff, CF14 4XN, UK (e-mail: marshallc3@cardiff.ac.uk)

S. Paisey is Pre-Clinical Facilities Manager, Wales Research and Diagnostic PET Imaging Centre, Cardiff, CF14 4XN, UK (e-mail: paisseysj@cardiff.ac.uk)

E. Spezi is a Reader at School of Engineering, Cardiff University, Cardiff, CF24 3AA, UK (e-mail: espezi@cardiff.ac.uk)

II. METHODS

The Mediso Nanoscan PET/CT was used to image eight mice with 4T1 tumours (mammography carcinoma xenografts) injected with 10.0 ± 2.0 MBq of 18F-FDG. They were imaged 50 minutes post injection for a duration of 20 minutes. Images were re-binned into four PET acquisitions each with a 5-minute duration corresponding to 50-55, 55-60, 60-65 and 65-70 minutes post injection. The first time point (50-55 minutes) was used for defining the segmentation contours. Four different systematic 3D-Contour sizes (4, 4.5, 5, 5.5 mm) were depicted using Velocity 3.2.1 software (Varian Medical Systems, Palo Alto, CA). Figure 1 shows the coronal and sagittal images with four contours on the lower flank of the first mouse. Contours defined on the first time point were used on all other images which were re-binned into subsequent time points SPAARC (Spaarc Pipeline for Automated Analysis and Radiomic Computing an in-house developed tool built on Matlab) was used to extract 78 3D-radiomic features for each of the four segmented volumes for each of the four time points [13]. Features including gray level co-occurrence matrix (GLCM), gray-level run-length matrix (GLRLM), gray-level size zone matrix (GLSZM), Gray-level distance zone matrix (GLDZM) and neighborhood gray-tone difference matrix (NGTDM) were extracted. All radiomics features were compliant with the International Biomarker Standardization Initiative (IBSI) [14].

The correlation between a variable n and p is defined using equations 1 and 2 (for contour size and imaging time respectively) [15]. To fully assess the statistical strength of correlation for each texture feature versus contour size four correlation matrices were constructed at each time point (T1 to T4) (C-T1, C-T2, C-T3, C-T4) using all contour sizes (C1 to C4). Similarly, this was repeated for probing the correlation of texture features versus acquisition time (T-C1, T-C2, T-C3, T-C4) giving four correlation matrices where time is the variable under investigation. The correlation of 78 features with varying time and contour size was probed resulting in 8 correlation matrices for each feature.

$$C_{n,p} = \frac{\sum_{i=1}^{m=8} (C_n - \bar{C}_n)(C_p - \bar{C}_p)}{(\sigma_{C_n} * \sigma_{C_p})(m-1)} \quad (1)$$

Similarly,

$$T_{n,p} = \frac{\sum_{i=1}^{m=8} (T_n - \bar{T}_n)(T_p - \bar{T}_p)}{(\sigma_{T_n} * \sigma_{T_p})(m-1)} \quad (2)$$

Were,

$C_{n,p}$: The correlation coefficient for n^{th} contour size and p^{th} contour size.

$\bar{C}_{n/p}$: The mean of all features value for n/p^{th} contour size.

$T_{n/p}$: feature value for n/p^{th} acquisition time.

$\sigma_{C_{n/p}}$: The standard deviation of all the features values for n/p^{th} contour size.

$T_{n,p}$: the correlation coefficient for n^{th} acquisition time and p^{th} acquisition time

$C_{n/p}$: feature value for n/p^{th} contour size.

$\bar{T}_{n/p}$: The mean of all features value for n/p^{th} acquisition time.

$\sigma_{T_{n/p}}$: The standard deviation of all the features values for n/p^{th} acquisition time.

Table I shows the resulting matrices for examination of robustness of texture features with varying contour sizes and imaging acquisition time. The determinant of the correlation matrix allows investigation into the linear dependence (or lack thereof) of the variable of interest (acquisition time or contour size) on the texture feature of interest.

TABLE I. ILLUSTRATIVE MATRICES OF RESULTING CORRELATION MATRICES FOR EACH FEATURE

<i>Variable of interest is contour size for $t=1,2,3,4$</i>				<i>Variable of interest is acquisition time for $c=1,2,3,4$</i>			
$C_{1,1}$	$C_{2,1}$	$C_{3,1}$	$C_{4,1}$	$T_{1,1}$	$T_{2,1}$	$T_{3,1}$	$T_{4,1}$
$C_{2,1}$	$C_{2,2}$	$C_{3,2}$	$C_{4,2}$	$T_{2,1}$	$T_{2,2}$	$T_{3,2}$	$T_{4,2}$
$C_{3,1}$	$C_{3,2}$	$C_{3,3}$	$C_{4,3}$	$T_{3,1}$	$T_{3,2}$	$T_{3,3}$	$T_{4,3}$
$C_{4,1}$	$C_{4,2}$	$C_{4,3}$	$C_{4,4}$	$T_{4,1}$	$T_{4,2}$	$T_{4,3}$	$T_{4,4}$

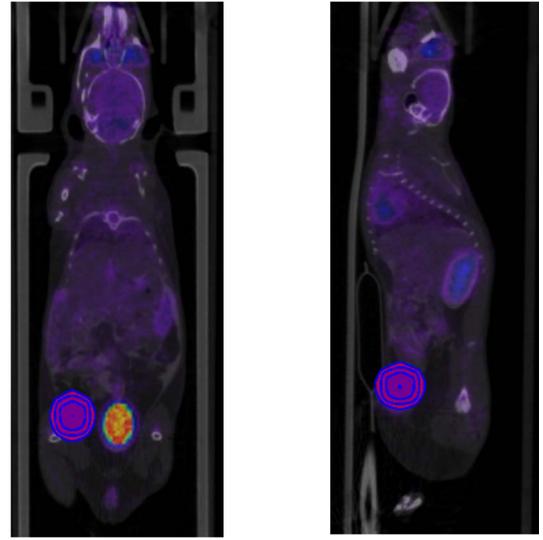


Fig. 1. Coronal (left) and sagittal (Right) slices of lower right flank (left of image) with four different contours for the first mouse.

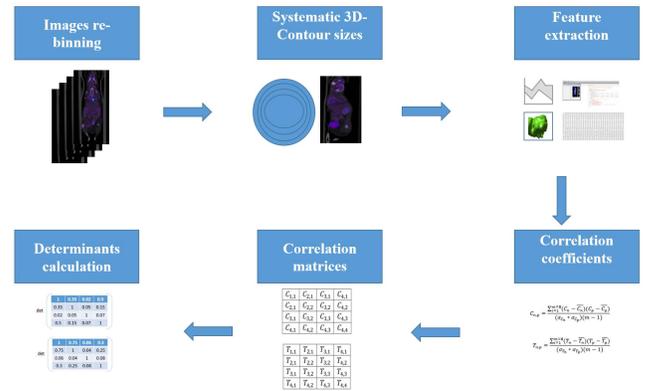


Fig. 2. Workflow for evaluating the statistical association of texture feature values versus different contour size and time points. Following acquisition images are re-binned into time points 50-55, 55-60, 60-65 and 65-70 minutes post injection. Each image has a 5 minute acquisition period. Four 3D contours were extracted at the 55 minute baseline scan and propagated to subsequent

time point images. 78 texture features were extracted for each image. For each feature 8 correlation matrices were constructed considering the correlation of a contour size and b) image time. The determinant of those matrices allows the robustness of each texture feature to contour size and acquisition time to be determined.

III. RESULTS

Overall, as shown in figure 3, there is a very small determinant of correlation when contour size is considered as the variable of interest. This is consistent when varying acquisition time for the majority of texture features. The mean determinant of correlation across all texture parameters when considering varying contour size is 0.02378. This suggests a linear dependence of texture features with increasing contour size. Less correlation and hence linear dependence is observed when acquisition time is considered as the variable of interest with a mean determinant of correlation across all texture parameters of 0.13296. Increasing contour size however has a variable effect on this metric of correlation.

Features including autocorrelation, information correlation 1, information correlation 2 (GLCM) and coarseness (NGTDM) were found to have 0 value as a mean determinant of correlation when considering varying acquisition time. In addition, there were no GLRLM, GLSZM and GLDZM features that have 0 value as a mean determinant of correlation for both contour and acquisition time. Table II and III show the radiomic features with highest and lowest values of mean determinant of correlation for acquisition time and contour size for each feature families.

Our result showed that 30 features, including; Joint entropy, Sum entropy, Dissimilarity (GLCM), Short runs emphasis (GLRLM), Strength (GNGTDM) have an increasing determinant of correlation with acquisition time as the contour size increased. Thirteen features including Run entropy (GLRLM) and Sum variance (GLCM), demonstrated a decrease in the determinant of correlation with time when the region of interest volume increased.

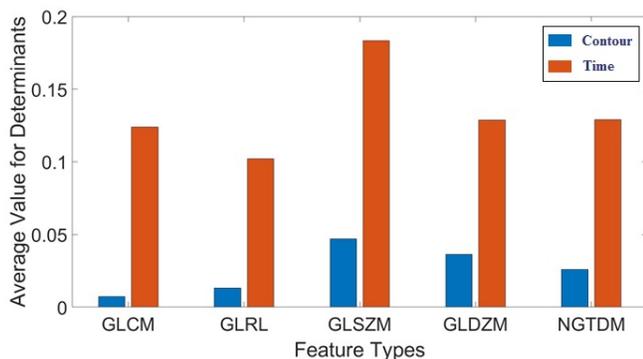


Fig. 3. Bar chart displaying the mean determinant of correlation for different features types whilst varying acquisition time and contour size.

The mean determinant of correlation across all GLCM texture parameters when considering varying contour size and acquisition time are 0.00733 and 0.12404, respectively. GLSZM features showed to have the highest mean

determinants of correlation at 0.18331 for acquisition time and 0.04688 for contour size.

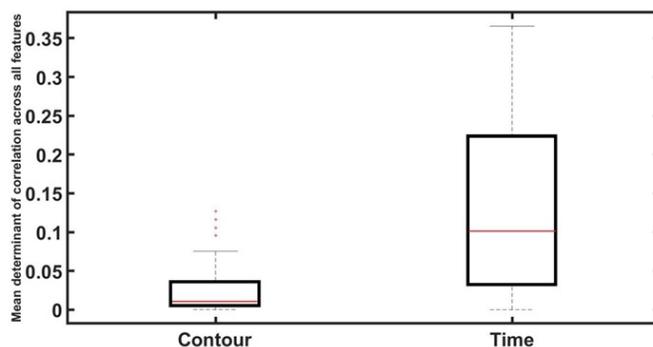


Fig. 4. Boxplot showing mean determinants of correlation matrices for all features whilst varying acquisition time (C-T1, C-T2, C-T3, C-T4) and contour size (T-C1, T-C2, T-C3, T-C4).

TABLE II. FEATURES WITH HIGHEST VALUE OF MEAN DETERMINANT OF CORRELATION FOR ACQUISITION TIME AND CONTOUR SIZE FOR EACH FEATURE FAMILIES.

	Contour Size - highest	Acquisition Time - highest
GLCM	Cluster shade	Difference variance
GLRL	Grey level non-uniformity	Short run high grey level emphasis
GLSZM	Zone size non-uniformity normalised	Small zone high grey level emphasis
GLDZM	Zone distance non-uniformity normalised	Grey level variance
NGTDM	Busyness	Contrast

TABLE III. FEATURES WITH LOWEST VALUE OF MEAN DETERMINANT OF CORRELATION FOR ACQUISITION TIME AND CONTOUR SIZE FOR EACH FEATURE FAMILIES.

	Contour Size - lowest	Acquisition Time - lowest
GLCM	Information correlation 1	AutoCorrelation, Information correlation 1, Information correlation 2
GLRL	Long run low grey level emphasis	Long run low grey level emphasis
GLSZM	Large zone low grey level emphasis	Large zone low grey level emphasis
GLDZM	Zone distance non-uniformity	Large distance emphasis
NGTDM	Coarseness	Coarseness

IV. CONCLUSIONS

The results demonstrate two things. First, texture features are more correlated with segmentation volume than acquisition time (mean determinant of correlation for segmentation volume = 0.02378 vs. mean determinant of correlation for acquisition time = 0.13296) and hence less robust to varying acquisition

time as showing in figure 4. Secondly the linear dependence / correlation of texture features with acquisition time is affected by contour size. This highlights that when performing radiomic analysis, image segmentation volume is correlated with the texture feature. A linear dependence may be enough to determine absolute margins of error on individual texture features with varying segmentation accuracy which provides scope for future work. Acquisition time however has a variable influence on this correlation and judicious choice of segmentation is required to minimize its effect. An implication of these findings is that both the impact of acquisition time and segmentation volume should be taken into account before radiomic analysis can be applied clinically.

REFERENCES

- [1] A. Rahmim and R. Wahl, "An Overview of Clinical PET/CT," *Iran. J. Nucl. Med.*, vol. 14, no. 26, pp. 1–14, 2006.
- [2] G. J. R. Cook, M. Siddique, B. P. Taylor, C. Yip, S. Chicklore, and V. Goh, "Radiomics in PET: principles and applications," *Neuroimage*, pp. 269–276, 2014.
- [3] R. L. Smith *et al.*, "Deep Learning Pre-Clinical Medical Image Segmentation for Automated Organ-Wise Delineation of PET," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 45, no. 1, p. S290, 2018.
- [4] R. L. Smith *et al.*, "Reinforcement Learning for Automated PET Image Segmentation," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 46, no. 1, p. S758-S759, 2019.
- [5] R. Forghani, P. Savadjiev, A. Chatterjee, N. Muthukrishnan, C. Reinhold, and B. Forghani, "Radiomics and Artificial Intelligence for Biomarker and Prediction Model Development in Oncology," *Comput. Struct. Biotechnol. J.*, vol. 17, pp. 995–1008, 2019.
- [6] R. T. H. Leijenaar *et al.*, "Stability of FDG-PET Radiomics features: An integrated analysis of test-retest and inter-observer variability," *Acta Oncol. (Madr.)*, vol. 52, no. 7, pp. 1391–1397, 2013.
- [7] I. El Naqa *et al.*, "Exploring feature-based approaches in PET images for predicting cancer treatment outcomes," *Pattern Recogn.*, vol. 42, no. 6, pp. 1162–1171, 2009.
- [8] G. J. R. Cook, G. Azad, K. Owczarczyk, M. Siddique, and V. Goh, "Challenges and Promises of PET Radiomics," *Int. J. Radiat. Oncol. Biol. Phys.*, pp. 1–7, 2018.
- [9] P. Galavis, C. Hollensen, N. Jallow, and B. P. Al., "Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters," *Acta Oncol. (Madr.)*, vol. 49, no. 7, pp. 12–22, 2010.
- [10] J. A. Oliver, M. Budzevich, G. G. Zhang, T. J. Dilling, K. Latifi, and E. G. Moros, "Variability of image features computed from conventional and respiratory-gated PET/CT images of lung cancer," *Transl. Oncol.*, vol. 8, no. 6, pp. 524–534, 2015.
- [11] E. Alsayed, R. Smith, C. Marshall, E. Spezi, and S. Paisey, "Stability of PET Radiomic Features: A Preclinical Study," in *European Journal of Nuclear Medicine and Molecular Imaging*, 2019, p. S759.
- [12] E. Pfaehler *et al.*, "Repeatability of 18F-FDG PET radiomic features: A phantom study to explore sensitivity to image reconstruction settings, noise, and delineation method," *Med. Phys.*, vol. 46, no. 2, pp. 665–678, 2019.
- [13] P. Whybra, C. Parkinson, and K. Foley, "Assessing radiomic feature robustness to interpolation in F-FDG PET imaging," *Sci. Rep.*, no. December, pp. 0–10, 2019.
- [14] A. Zwanenburg, M. Valli, and L. Steffen, "image biomarker standardisation initiative," vol. 1, 2016.
- [15] A. G. Asuero, A. Sayago, and A. G. Gonz, "The Correlation Coefficient: An Overview The Correlation Coefficient: An Overview," *Crit. Rev. Anal. Chem.*, vol. 36, no. January, 2006.