Computer aided quantification of prostate cancer diffusion-weighted imaging: repeatability analysis of radiomics as biomarkers for Gleason score prediction

Ileana Montoya Perez[1,3,6], Jussi Toivonen[1,3,6] , Parisa Movahedi[1,3,6], Harri Merisaari[1,2,6], Janne Verho[1], Pekka Taimen[4], Peter J. Boström[5], Hannu J. Aronen[1,6], Tapio Pahikkala[3], Ivan Jambor[1,2,6]

[1] Department of Diagnostic Radiology, University of Turku, Turku, Finland
[2] Turku PET Centre, University of Turku, Turku, Finland
[3] Department of Future Technologies, University of Turku, Turku, Finland
[4] Department of Pathology, University of Turku and Turku University Hospital, Turku,Finland
[5] Department of Urology, Turku University Hospital, Turku, Finland
[6] Medical Imaging Centre of Southwest Finland, Turku University Hospital, Turku, Finland

Sypnosis

We evaluated the repeatability of apparent diffusion confident, derived using monoexponential function (ADCm) from prostate cancer DWI (12 b values, 0-2000 s/mm^2), radiomics of prostate cancer and their potential to predict prostate cancer Gleason score (histological grading system of prostate cancer aggressiveness). Statistical features (mean, median, 10th, 25th percentile) and Gabor texture feature of DWI ADCm parametric maps showed high repeatability and correlated significantly with Gleason score. In contrast, homogeneity gray-level co-occurrence matrix showed low repeatability despite having significant correlation with Gleason score.

Introduction

Prostate Cancer (PCa) is one of the most common cancers among men in United States.[1] Although the number of PCa diagnosed cases is high, the proportion of PCa related deaths is substantially lower. The majority of men diagnosed with PCa have a non-aggressive tumor that rarely develops into a deathly disease.

The aggressiveness of the tumor is commonly graded using a Gleason system.[2] This system is used for biopsy and prostatectomy specimens. Correctly classifying tumors based on their Gleason score helps in planning patient treatment. However, Gleason score obtained from transrectal ultrasound-guided biopsies (TRUS) is inaccurate because of sampling error. Therefore, additional ways of predicting the tumor Gleason score are desired.

Diffusion weighted imaging (DWI) have shown potential for non-invasive detection and characterization of PCa[3–6]. It has been suggested that radiomics derived from apparent diffusion confident, derived using monoexponential function (ADCm), can have a role in Gleason score classification[8]. However, repeatability of these approaches has not been evaluated. In this study, we assess the repeatability of ADCm statistical and texture features and their potential to predict PCa Gleason score.

Materials and Methods

The study was approved by the institutional review board and all patients provided written informed consent. A total of 55 patients with histologically confirmed PCa, scheduled for prostatectomy, were prospectively enrolled and underwent two prostate MR examinations performed on the same day. The DWI data sets were acquired using a 3T MR scanner (Ingenuity PET/MR, Philips, Cleveland, USA), SE-EPI sequence, monopolar diffusion gradient scheme, gradient overplus option on and the following parameters: TR/TE 3141 ms/51 ms, FOV 250×250 mm2, acquisition matrix 100x99, reconstruction matrix 224x224, slice thickness 5 mm, SENSE factor of 2, 12 b values of (number of signal averages) 0 (2), 100 (2), 300 (2), 500 (2) (2), 700 (2), 900 (2), 1100 (2), 1300 (2), 1500 (2), 1700 (3), 1900 (4), 2000 (4) s/mm2; diffusion time of 20.3 ms. The acquisition time was 8 min and 29 s. PCa lesions were delineated using whole mount prostatectomy sections as the "gold standard".

The first-order statistical features were obtained from the whole lesion ADCm histograms, these included mean, median, 10th and 25th percentile. Gabor texture was extracted using squared-shaped window of evenly spaced voxels (side length 11), frequency for the sinusoidal ($f = 0.3$), a circular Gaussian envelop ($\sigma = 1$), and the mean of the real part. Homogeneity gray-level co-occurrence matrix (H-GLCM) feature was extracted using a minimum bonding box around the whole lesion on each slice, non-lesion voxels were excluded.

The repeatability of the features was assessed using within-subject variance (σ_w^2), intraclass correlation coefficient (ICC) – specifically ICC(3,1)[9]–, repeatability coefficient (RC) and within-subject coefficient of variation (wCV, %). Correlation of the features with Gleason score groups (3+3, 3+4, >3+4) was evaluated using Spearman correlation coefficient (ρ).

Results

In total 47 patients (47/55, 86%) and 71 tumors were included in final analyses.

Shapiro-Wilk normality test showed that the statistical features, Gabor, and a logarithm transformation of H-GLCM followed a normal distribution. Whole mount histopathologic sections (RALP), ADCm and texture maps of three patients are presented in Figure 1. A comparison of the ADCm features for all lesions is shown in Figure 2. All statistical features showed high repeatability with ICC values raging from 0.79 to 0.85, σ_w^2 from 3.9e-09 to 5.4e-09, RC from 1.7e-04 to 2.0e-04, and wCV from 6% to 9%. Gabor feature demonstrated high repeatability with ICC value of 0.80 (0.70 – 0.87), σ_w^2 =4.4e-09, RC=1.8e-04 and wCV of 9% (Table 1). H-GLCM showed poor repeatability with ICC values of 0.35 (0.13 – 0.54) although statistically significant ($p<0.0001$) ρ value of -0.50. Gabor texture feature had the highest ρ value of -0.59 ($p<0.0001$) while the statistical features had ρ values in the range of -0.31 to -0.43 (Table 2).

Discussion

To the best of our knowledge this is the first study evaluating repeatability of PCa DWI radiomics. We have shown that statistical features and Gabor texture feature have high

repeatability and correlate significantly with PCa Gleason score. However, H-GLCM had significant correlation with PCa but failed to demonstrate high repeatability which is crucial in patient management of patients with clinically suspected or histologically confirmed PCa.

Conclusion

Statistical features (mean, median, 10^{th}, 25^{th} percentile) and Gabor texture of DWI ADCm parameteric maps showed high repeatability and correlated significantly with Gleason score groups. In contrast, H-GLCM showed low repeatability. Our analyses suggest that Gabor texture feature has a potential to improve quantification of prostate DWI by providing high repeatability and improved Gleason score estimation.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017. *CA Cancer J Clin*. 2017;67(1):7-30. doi:10.3322/caac.21387.
2. Epstein JI, Allsbrook WCJ, Amin MB, Egevad LL. The 2005 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma. *Am J Surg Pathol*. 2005;29(9):1228-1242. doi:10.1097/01.pas.0000173646.99337.b1.
3. Quentin M, Blondin D, Klasen J, et al. Comparison of different mathematical models of diffusion-weighted prostate MR imaging. *Magn Reson Imaging*. 2012;30(10):1468-1474. doi:10.1016/j.mri.2012.04.025.
4. Toivonen J, Merisaari H, Pesola M, et al. Mathematical models for diffusion-weighted imaging of prostate cancer using b values up to 2000 s/mm2: Correlation with Gleason score and repeatability of region of interest analysis. *Magn Reson Med*. 2015;74(4):1116-1124. doi:10.1002/mrm.25482.
5. Boesen L, Chabanova E, Logager V, Balslev I, Thomsen HS. Apparent diffusion coefficient ratio correlates significantly with prostate cancer gleason score at final pathology. *J Magn Reson Imaging*. 2015;42(2):446-453. doi:10.1002/jmri.24801.
6. Donati OF, Afaq A, Vargas HA, et al. Prostate MRI: Evaluating tumor volume and apparent diffusion coefficient as surrogate biomarkers for predicting tumor Gleason score. *Clin Cancer Res*. 2014;20(14):3705-3711. doi:10.1158/1078-0432.CCR-14-0044.
7. Donati OF, Mazaheri Y, Afaq A, et al. Prostate cancer aggressiveness: assessment with whole-lesion histogram analysis of the apparent diffusion coefficient. *Radiology*. 2014;271(1):143-152. doi:10.1148/radiol.13130973.
8. Fehr D, Veeraraghavan H, Wibmer A, et al. Automatic classification of prostate cancer Gleason scores from multiparametric magnetic resonance images. *Proc Natl Acad Sci U S A*. 2015;112(46):E6265-73. doi:10.1073/pnas.1505935112.
9. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86(2):420.

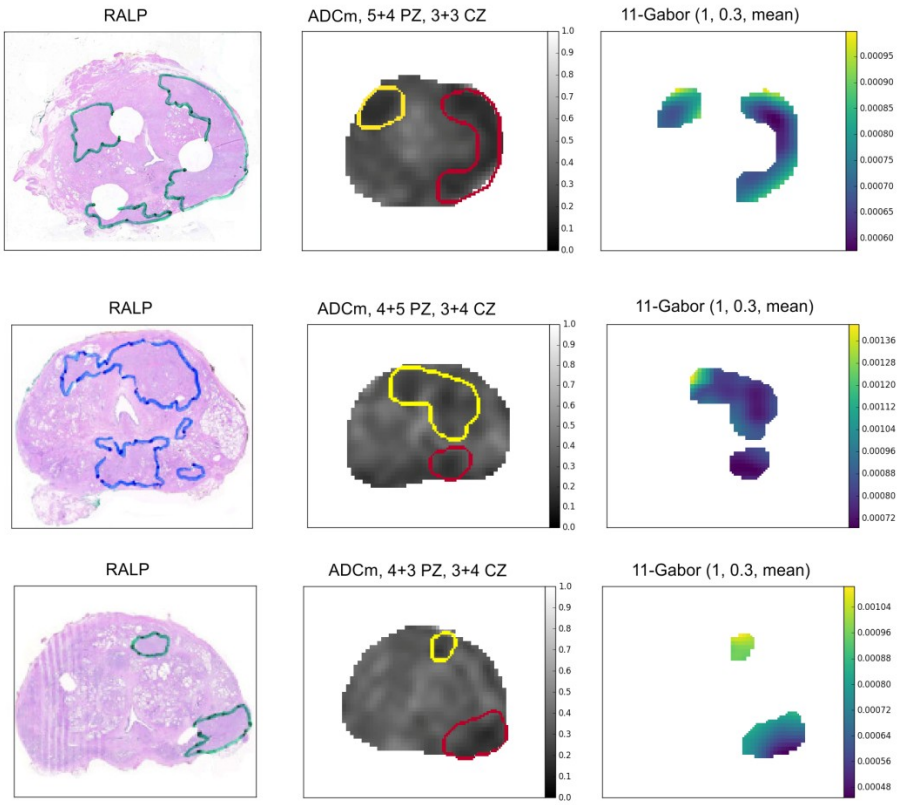

Figure 1: Three representative cases with corresponding histopathology slice (RALP), ADCm map and Gabor texture map.

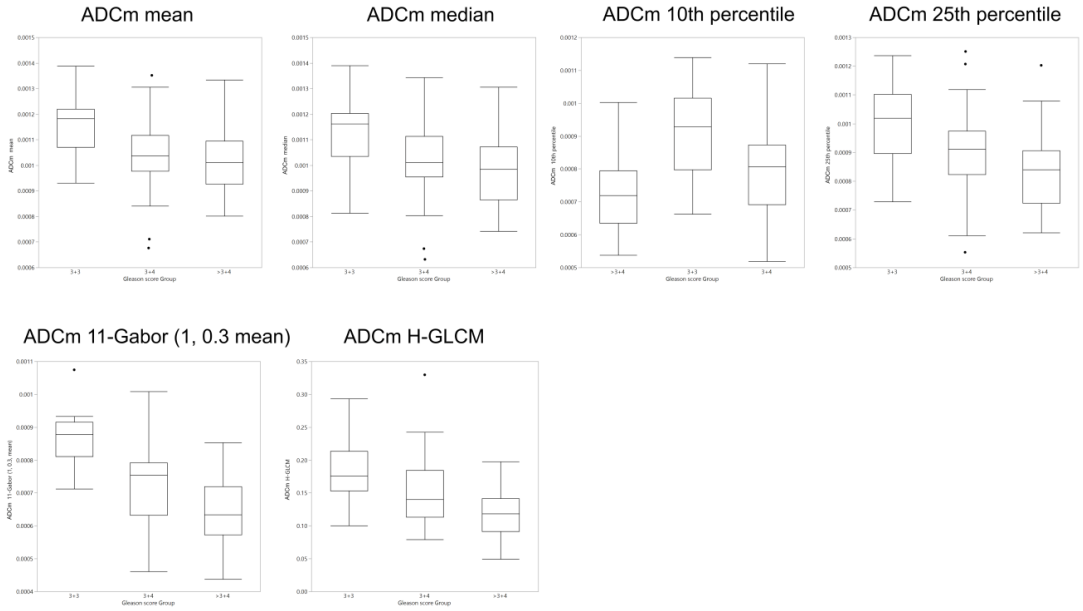

Figure 2: Boxplot show comparison of ADCm radiomics for all lesions. Statistical features (upper row) and texture features (lower row). Line in box is median, height of the box represents interquartile range, and points indicate outliers.

Table 1: Repeatability analysis for ADCm radiomics across two repeated scans. Intraclass correlation coefficient (ICC), within-subject variance (σ_w^2), repeatability coefficient (RC) and within-subject coefficient of variation (wCV, %).

Feature	**ICC (95% CI)**	**σ_w^2**	**RC**	**wCV**
Mean	0.80 (0.69, 0.87)	4.94e-09	0.000195	0.07
Median	0.85 (0.77, 0.90)	4.03e-09	0.000176	0.06
10^{th} Percentile	0.79 (0.69, 0.87)	5.46e-09	0.000205	0.09
25^{th} Percentile	0.85 (0.77, 0.91)	3.96e-09	0.000174	0.07
11-Gabor (1,0.3,mean)	0.80 (0.70, 0.87)	4.39e-09	0.000184	0.09
H-GLCM	0.35 (0.13, 0.54)	1.23e-01	0.970587	0.36

Table 2: Spearman correlation coefficient (ρ) for correlation analysis between ADCm features and Gleason score. * significance with p<0.0001

Feature	**ρ**	**p-value**
Mean	-0.33	0.0047
Median	-0.31	0.0092
10^{th} Percentile	-0.43	0.0002
25^{th} Percentile	-0.40	0.0005
11-gabor(1,0.3,mean)	-0.59	<.0001*
H-GLCM	-0.50	<.0001*