MR. CLEMENT ORCZYK (Orcid ID : 0000-0003-3067-2409)

PROF. MARK EMBERTON (Orcid ID : 0000-0003-4230-0338)

Article type : Original Article

Article category: Urological Oncology

Prostate cancer heterogeneity: texture analysis score based on multiple MRI sequences for detection, stratification and selection of lesions at time of biopsy

Authors: Clement Orczyk[1,2,3,8],MD,PhD, Arnauld Villers[4], MD,PhD, Henry Rusinek[5],PhD, Vincent Lepennec[6],MD, Céline Bazille[7],MD,PhD, Francesco Giganti[1,9],MD, Artem Mikheev[5],MSc, Myriam Bernaudin[3],PhD, Mark Emberton[1,2],MD, Audrey Fohlen[3,6],MD, Samuel Valable[3] PhD

[1]Division of Surgery and Interventional Sciences, University College London, London, United Kingdom

[2] Department of Urology, University College London Hospitals, London, United Kingdom

[3] Normandie Univ, UNICAEN, CEA, CNRS, ISTCT/ CERVOxy group, GIP CYCERON, 14000 Caen, France

[4]Department of Urology, University Hospital of Lille, Nord de France

[5]Department of Radiology, New York University Medical Center, New York, NY, USA

[6]Department of Radiology, University Hospital of Caen, Caen, France

[7]Department of Pathology, University Hospital of Caen, Caen, France

[8]Department of Urology, University Hospital of Caen, Caen, France

[9]Department of Radiology, University College London Hospital, London, UK

Corresponding author: Mr Clement Orczyk, MD, PhD

Associate Professor of Urology, Honorary Consultant Urological Surgeon

UCL, Charles Bell House

43-45 Foley St

W1W 7TS, London UK

clementorczyk@yahoo.fr

c.orczyk@ucl.ac.uk

mobile: +44 7 958 550 727

fax: +44 2 07679 9511

Key word: prostate cancer, detection, stratification, MRI, texture analysis, histogram, pharmacokinetic model, image processing, biopsy, entropy, radionomics

Abstract

Purpose: To undertake an early proof of concept study on a novel, semi-automated texture-based scoring system in order to enhance the association between MRI lesions and clinically significant cancer.

Patients and methods: With ethics approval, 536 imaging volumes were generated from 20 consecutive patients who underwent mpMRI at time of biopsy. Volumes of interest (VOIs) included zonal anatomy segmentation and suspicious MRI lesion for cancer (Likert scale score greater than 2). Entropy (E), measuring heterogeneity, was computed from VOIs and plotted as a multiparametric score defined as Entropy Score (ES) = E ADC+ E Ktrans + E Ve+ E T2WI. The reference test that was used to define the ground truth comprised systematic saturation biopsies coupled with MRI targeted sampling. This generated 422 cores in all that were individually labelled and oriented in 3D.

Diagnostic accuracy for detection of clinically significant prostate cancer (SPCa), defined as Gleason score of 3+4 (or higher) or more than 3mm of any grade of cancer on a single core, was assessed using Receiver Operating Characteristics, correlation and descriptive statistics. Proportion of cancerous lesions detected by ES and Visual Scoring (VS) were statistically compared using paired McNemar test.

Results: Any cancer (Gleason Score 6 to 8) was found in 12 of the 20 (60%) patients with a median PSA of 8.22ng/ml. SPCa (ES=17.96 ±0.72 NAT; CI 95%) showed a significant higher ES than non-SPCa (ES=15.33 ±0.76 NAT). ES correlated with Gleason Score (rs =0.5683, p=0.033) and maximum cancer core length (ρ = 0.781; p=0.0009). The Area Under the Curve for ES (0.89) and visual scoring (VS) (0.91) were not significantly different (p=0.75) for detection of SPCa among MRI lesions. Best ES estimated numerical threshold of 16.61 NATural information unit (NAT) led to a sensitivity of 100% and negative predictive value of 100%. The proportion of MRI lesion which found to positive for SPCa using this ES threshold (54%) was significantly higher (p<0.001) than those using VS (24% of score 3,4,5) in a paired analysis using McNemar test. 53% of MRI lesion would have avoided biopsy sampling without missing significant disease.

Conclusion: Capturing heterogeneity of PCa across multiple MRI sequences with ES yielded high performances for the detection and stratification of SPca. ES outperformed visual scoring in predicting positivity of lesions, holding promise in the selection of targets for biopsy and calling for further understanding of this association.

Key words: prostate cancer, detection, stratification, MRI, texture analysis, histogram, pharmacokinetic model, image processing, biopsy, entropy, radionomics

Introduction

Current diagnostic strategies for prostate cancer detection have relied upon prostate specific antigen (PSA) and random sampling of the gland using transrectal ultrasound (TRUS) guided biopsy. We now know that this approach is associated with an over-representation of clinically insignificant cancer as well as an under-representation of clinically significant cancer (1).

MpMRI has been introduced into the pathway in order to mitigate some of the deficiencies associated with the standard of care. Its introduction has been associated with an increased detection of clinically significant PCa and, in certain circumstances it has provided a strategy to address the problem of over-diagnosis(1–3) raised in major screening trials(4–6).

In current practice, prostate mpMRI visual reporting requires considerable expertise and is prone to inter-observer variability even when applying standardised guidelines, like Prostate Imaging Reporting And Data System (PI-RADS) version 2.0 (7). It also takes a considerable amount of time to do given the number of sequences that need to be looked at and compared with each other and possibly with previous imaging. Quantitative parameters derived from functional sequences, do not offer a reliable alternative as they are prone to substantial variabilities (8).

Some recent developments in image analysis (by means of neural networks) are now making the possibility of automatic MRI-segmentation of both the prostate itself and of the MRI-lesions within it (9,10). Automated assessment of the lesions, if it were indeed possible, opens up the possibility of fully automated MRI reporting in the future and access to and computation of information not perceived by visual assessment in the context of radiomic analysis of standard of care images (11,12).

Tumour heterogeneity is considered to be one of the main drivers of progression and resistance to treatment (13). The degree of histological PCa heterogeneity (i.e. the gradual loss of a structured cellular architecture) is reflected at tissue level by the Gleason grade (14,15) and was found at an intratumoral level as opposed to benign tissue (15–17) by whole genomic sequencing.

Heterogeneity of PCa may also be noninvasively assessed using parameters derived from imaging. Entropy computes heterogeneity of a given volume (12,18,19), measuring spatial randomness of image intensities. A higher entropy represents a more disordered distribution of values within the tissue (20), potentially reflecting biological feature of heterogeneity of prostate cancer, as in other carcinomas using MRI (20).

There are evidences suggesting that lesion entropy could be derived from specific mpMRI sequences for the detection, stratification and prognosis of PCa (21–24). Combining Entropy from different MRI sequences might improve its ability for cancer detection and stratification, as image generated by each sequence reflects different aspects of the underlying lesion histopathology and does not necessary detect a cancer focus due to histology variation in tissue composition(25).

Our aim is to undertake an early proof of concept study on a novel texture-based scoring system, relying on entropy, in order to enhance the association of MRI derived target generation with clinically significant cancer.

Material and Methods

Study population and image acquisition (standard test)

Under local ethics committee approval, that waive written informed consent, we analysed the data of 21 consecutive patients referred for suspicion of PCa without contraindications to MRI. Patients underwent mpMRI prior to biopsy except those with suspicious of high risk disease. In such case, biopsies were carried out first. All patients were scanned on a 1.5 T MRI system with a pelvic phased-array (Siemens Magnetom Avanto), using the same mpMRI protocol (T2 weighted imaging, DWI with generation of ADC maps and DCE imaging) in 2013 and 2014. The details of the protocol are listed in Appendix Table 1 and are in line with the European Society of Urogenital Radiology (ESUR) 2012 guidelines that were available at the start of the study (26).

Two radiologists (with 6 and 3 years experiences in prostate mpMRI interpretation respectively) reported the scans according to a Likert scale from 1 to 5 (27). All sequences were used to allocate the score. Lesions with scores 3 to 5 were considered to be suspicious for cancer. After consensus review of the images with the urologist performing the biopsy, a map of 36 regions of interest (ROIs) was generated before the procedure(28). Visual scoring of MRI was the standard test.

Reference standard: Biopsy procedure and correlation

All patients underwent subsequent transrectal ultrasound (TRUS) guided biopsy as per local protocol which implied a saturation TRUS biopsy for PSA under 15ng/ml with a 22 cores template. Each core, either systematic or targeted, was labelled according to its location and sent in a separate jar for analysis. In case of suspicious lesion at mpMRI prior biopsy,

targeted biopsies were performed using cognitive registration by a single urologist trained to this technique, prior systematic sampling(29). A dedicated uro-pathologist reported the biopsy. A drawing reporting location and numbering of the cores was issued for each patient and sent to pathology. Systematic TRUS biopsy combined with targeted biopsy corresponds to the reference test. Significant disease was defined as the presence of Gleason pattern 4 (primary or secondary) and/or more than 3 mm cancer in a single core (30).

Biopsy and MRI maps were compared in consensus to establish concordance for cancer leading to true positive, false positive, false negative and true negative MRI regions.

Index test : Image processing

Image processing protocol is illustrated in Appendix figure 1.

Quantitative maps from DCE imaging were computed with correction of motion artefacts. The two parameters were the Volume Transfer Constant (Ktrans in ml/min) and Fraction of extracellular extravascular space (Ve, no unit) (see Appendix 1 for processing pharmakocinetic model).

All quantitative maps, ADC map and 3D DCE-derived maps (Ktrans and Ve), were aligned in the anatomical reference space (T2-WI) using automatic rigid registration. Therefore, the same volume was analysed across the difference sequences.

Masks to generate Volume of interest (VOI) were manually segmented in T2-WI within a dedicated platform by an individual trained across datasets of prostate mpMRI-histology correlation and performing image guided procedure based on mpMRI. The mask segmentation was then propagated across sequences. The different VOIs according to

zonal anatomy were the prostate (excluding seminal vesicles), the peripheral zone (PZ), the transition zone (TZ). Patient related VOIs were the MRI suspicious lesions at Visual Scoring, normal peripheral zone (nPZ) and benign prostate hyperplasia nodules (BPHn). Those are referred as discrete VOIs.

As results, a set of VOIs at MRI was generated for: prostate, PZ, TZ, nPZ, BPHn, non-cancerous and cancerous area, including lesion at Visual Scoring for 4 imaging parameters.

Texture analysis was computed from the VOIs as a 3D assessment using a statistical analysis of the histogram, known as 1st order analysis (11,18,22). Entropy was directly computed for the VOI, corresponding to the whole lesion in 3D (19) (appendix 3).

The computation entropy has been calibrated to sample the relevant information from a 0.2cc lesion, as considered as the minimum volume of significant disease when Gleason pattern 4 is identified (31). Methods of calibration are detailed in Appendix 3. Consequently, all sequences are sampled with the same precision for the relevant information.

Entropy unit is NATural unit of information (NAT) for each parameter.

To assess the entropy derived from different images, we introduced an entropy score (ES) representing the combined entropy within each VOI to characterise quantitatively a volume in the same manner mpMRI is visually reported with a multiparametric approach.

ES for a v is defined in the present work as

$$ES(v) = E^{ADC}(v) + E^{Ktrans}(v) + E^{Ve}(v) + E^{T2WI}(v)$$

Statistical analysis:

Power calculation was performed to demonstrate significant differences in the mean value of Ktrans between cancer foci and benign tissue (($\alpha = 0.05;\ \beta = 1 - 0.9$), two-tailed hypothesis).

Statistical analysis was performed by using R software (The R Project for Statistical Computing, Vienna, Austria, http://www.R-project.org). Figure 1 describes the different level of analysis and study workflow.

The quantitative parameters and derivative metrics were tested as paired data to compare the values of cancer VOIs to the matching normal tissue (either nPZ or BPHn depending of the zone of origin of the tumour).

We used a variance analysis test of Friedman for paired data to assess if the tested parameters were independent for a given type of VOI.

Performances for detection were assessed at two levels of analysis (figure1). Analysis A considered all VOIs related to patient. Analysis B considered only visually scored lesions (either cancerous or not).

For continuous data and Entropy score at the two aforementioned level of analysis, Receiving Operator Characteristics (ROC) curves were generated and AUC calculated (*pROC* package) to perform binary classification between significant PCa (SPCa) and and non SPCa (either insignificant or non cancerous). Comparison of the AUC (significance if $p<0.05$) of ES was carried out against each ES component taken individually. To fulfil primary objective we compared AUC of Visual Scoring (VS) to Entropy Score (ES).

Logistic regression (LR) modelling was performed for both A and B analysis including each single entropy parameter and promising raw quantitative value for Ktrans, Ve and ADC and clinical data. Performances of the model were assessed using ROC curves for both A and B analysis. The AUC of the LR models, VS and ES were compared.

Sensibility, specificity, positive and negative predictive values are also calculated for LR models and the ES for both level of analysis using best calculated threshold from the ROC. Best estimated threshold was computed as maximizing the sum of Se and Sp.

Applying this best estimated threshold, we compared proportions of positive lesions for SPCa estimated by VS and ES in the B analysis population with a paired McNemar test.

The correlation of ES with Gleason score was performed using the Spearman coefficient. The Correlation of ES with the Maximum Core Length was assessed using the Pearson coefficient.

All statistical tests were conducted at the two-sided 5% significance.

Results:

Population and descriptive results:

From the 21 patients enrolled, one patient found to be ineligible for inadequate DCE imaging.

From the 20 remaining patients, PCa was detected among 12 (60%) patients following those biopsies either form systematic or targeted cores. Patient characteristics are shown in Table 1. Reference test represents a data set of 442 biopsy cores independently labelled and oriented in space. Median time between MRI and biopsy was 35 days. The total number of targets was 28 of which 7 were positive for cancer. Targets were scored 3, 4 and 5 in respectively 20, 5 and 3 cases. Detection rate of any cancer was 2/20 (10%), 2/5 (40%) and 3/3 (100%) for the same range of score. All targets from MRI were segmented in the post processing software.

Full imaging workflow was successfully carried out in all cases, leading to a data set of 536 independent volumes. After correlation with histology findings, 60 MRI sequences specific volumes were generated from 15 cancer foci after fusion of the sequences and maps in the T2WI space. Histogram generation and texture analysis were available for all the 536 VOIs.

The descriptive results of quantitative parameters, derived metrics and ES are plotted in table 2. Appendix Table 3 shows mean values for quantitative parameters and E in paired match analysis between cancer VOIs and normal paired tissue.

There was a consistent trend for Entropy of each parameter for positive targets to be higher than negative as shown in box plot in appendix figure 2.

The Friedman test showed independence (p=0.0005) of each individual component of the ES enabling a meaningful score.

For stratification of the disease, significant cancer (ES=17.73 ±0.1.23 NAT) showed a significant higher Entropy Score than non-significant cancer (ES=15.33 ±0.78 NAT, p=0.0025).

The Pearson‘s correlation coefficient of the ES of cancerous VOIs with the maximum cancer core length (MCL) found to be positive with ρ = 0.781 (p=0.0009).

The Spearman’s correlation coefficient of the ES of cancerous VOIs with matching biopsy Gleason score was also positive with r_s =0.5683 and p=0.033.

Performances of the Index test (ES) for discrimination of MRI lesions:

Table 3 summarizes diagnostic performances for ES and regression model for both A and B analyses.

For analysis A, there was a significant difference (p=0.0009) in ES between positive VOIs for significant cancer (mean 17.973± 01.23 NAT; CI 95%) and those negative (mean 15.33 ± 0.78 NAT).

ROC curves were built for the ES, entropy and mean value of each individual MRI parameter for detection of significant cancer. ROC curves were generated and AUC calculated from a total of 240 imaging volumes representing 70 VOIs from MRI suspicious lesions and normal tissue (figure 2). AUC results and comparison are summarized in appendix table 4.

ES achieved an AUC of 0.88 (0.76, 0.97; 95% CI) and was higher than any other single parameter. Best threshold for ES was 16.69 NAT with Sensitivity 100% (95% CI 63.06 – 100) of and Specificity of 69.35% (95% CI 56.35 - 80.44). The diagnostic accuracy was 72.86 (95% CI 60.9 - 82.8). The Youden's index was 0.69 (95% CI 0.19 - 0.80)

For analysis B, there was a significant difference (p=0.001) in ES between VOIs positive for significant cancer (mean 17.21± 1.00 NAT) and those negative (mean 16.11 ± 0.1.19 NAT; CI 95%).

The same ROC curves (figure 2) were built for the 132 imaging volumes representing the 33 MRI lesions (targets of score greater than 2 and cancer). The AUC for the ES to detection SPCa was 0.89 (0.76, 0.99; 95% CI), with a trend to be higher than other E or quantitative value for any other parameter (appendix table 4).

Best ES threshold was found for 16.61 NAT, with sensitivity of 100% and specificity of 72%.

Using this threshold of 16.61 NAT for detection of significant PCa, we calculated a negative predictive value (NPV) of 100%, Positive Predictive Value (PPV) of 53.3%. The diagnostic

accuracy was 78.79 % (95% CI 61.09 - 91.02). Youden index of ES was 0.72 (95% CI 0.14 - 0.88).

AUC of LR model using E Ktrans, E VE, E T2WI, mean ADC did not show significant difference to AUC of ES with value of 0.93 for both A and B analysis.

Performances of index against standard test: Entropy Score vs. Visual Scoring

ROC curve was plotted for VS in B analysis. AUC reached 0.91 which was not significant either with ES (p= 0.75) or Logistic regression model (p=0.74) as shown in figure 2.

18 MRI lesions over the 33 were under the ES threshold of 16.61 NAT from analysis B. None of them harboured significant PCa. (table 4).

The proportion of lesion which was found to be positive using this ES threshold (53%) was significantly higher ($p<0.001$) than positive lesions with VS (24% of score 3,4,5) in a paired analysis using McNemar test.

An estimated detection rate for score 3, 4 and 5 MRI lesions would have been respectively 52%, 100% and 100% and result in avoiding to sample 54% of the MRI lesions if the estimated ES threshold was applied in this data set.

Discussion

In this report, we focus on MRI generated lesion for detection and stratification of SPca. We used the concept of radiomics and texture analysis to enhance the association of imaging findings with SPCa. Particularity of this preliminary report relies on the choice of a feature, heterogeneity, that can be both identified in biology of cancer and imaging features. We showed an association of a texture feature, heterogeneity, with the presence of SPCa within an MRI lesion.

Our report illustrates the growing interest in the use of radiomics to mine the amount of information that medical imaging presents to increase the diagnostic accuracy(12,32,33). Also, the recent developments in the field of image analysis permit to envisage automatic segmentation of prostate MRI(9,34,35). This is therefore critical need to link imaging findings to biology and cancer outcomes(36).

In oncology, intra tumour heterogeneity drives neoplastic progression(13). Regarding PCa, recent reports further explored the heterogeneity of PCa at molecular and genetic analysis(17), even when considering the intratumoral level(15,16). It has been previously reported that heterogeneity of PCa is linked to prognosis(15).

MRI signature of prostate cancer is complex, heterogeneous and still not fully decrypted. Neovascularization, cellularity, proportion of histological components are reported to quantitatively and qualitatively impact overall detection of prostate cancer across different sequences (37,38). Even subtype of Gleason influences detectability of cancer foci by mpMRI (39).

We questioned in this experiment the association between this biological feature of heterogeneity and MRI signature of SPCa using texture analysis.

This study presents original aspects in the unified analysis of texture of three different sequences, incorporating T2-WI, DWI, DCE imaging, within one quantitative score capturing heterogeneity of distinct tissue component of PCa in the purpose of validation of MRI lesions for biopsy. Capturing radiomics features across multiple MRI sequences has already been described for breast, liver and nasopharingeal carcinomas and even linked to stratification and prognosis of the disease (20,40,41). To our knowledge no unified score based on radiomics features extracted from MRI findings has been described in this purpose. This work adds to the hypothesis of texture analysis as a link between biology of carcinomas and its visualization using MRI.

There was a significant difference in ES between SPCa and benign tissue and as well with insignificant cancer foci. Clinical potential relevance for stratification of the disease is further illustrated by the strong positive correlation with GS, important driver of oncologic outcome (42), and MCL. ES presented stronger correlation (r_s = 0.5683 ; p= 0.03) with GS than previous reports of inverse correlation of ADC of r=– 0.376 by Oto et Al.(43) or r=– 0,39 by Verma et al.(44). Entropy of ADC for a whole lesion was found to be an independent predictor of biochemical failure by Rosenkrantz et al(24)

Supporting the potential value of radiomics, Vignati et al.(45) found excellent AUC of 0.96 for a texture parameter based on T2WI alone for differentiating Gleason 6 and >6 in a radical prostatectomy cohort, limiting translation in a diagnostic setting.

In a recent report, Ginsburg et al.(46) trained a model based on numerous radiomics features in multi-institutional data-set for detection of cancer on a voxel basis. Their maximal AUC of 0.71 was lower than our results and based on a single sequence analysis. Rosenktrantz et al.(47) applied different features of 1st order to discriminate lesions with Gleason >6 to others among visually scored lesions. Harboring similar range of values to our results, Entropy of ADC alone failed to discriminate those even in their larger cohort.

In our experiment, the lack of significance of entropy of each sequence taken individually (figure 2) underlies that heterogeneity assessment by MRI should be

performed through a multiparametric approach, as diagnosis is performed with visual scoring(48,49).

As a multi-sequence strategy, ES achieved similar and very good performances to a logistic regression model integrating imaging features (including mono-sequential entropy), by comparison of AUC (0.93 vs 0 .89, p=0.32). ES outperformed the LR model when considering NPV (100%) and Se (100%). This is important findings to envisage generalization of ES. A LR model has to be calibrated for a given population when ES needs calibration of the sequences acquisition which can be achieved with phantoms (33). Pending validation studies, we could envisage development of pre-calibrated phantom to setup MRI scanners of different magnetic field strengths associated with automated reporting of MRI.

Visual Scoring already has been reported of utility for detection of PCa using on targeted approach in the Precision trial, as multicenter randomized control trial against random TRUS biopsy(2). Detection of significant disease was 12% higher with targeted biopsy compared to standard TRUS biopsy. Proportion of insignificant disease was also lower using the MRI guided strategy.

The Precision trial also successfully tested the strategy to avoid biopsy in 28 % patient with negative imaging findings(2). However, Visual scoring of mpMRI was described with variability, especially in NPV ranging from 63% to 95%(50). High and replicable negative predictive value would avoid to unnecessary sample some lesion without missing the clinically significant cancer while high sensitivity is conserved.

Process leading to visualisation of MRI lesions and radiologists' general impression appear to be difficult to systematize as expressed by non-negligible inter observer variability of visual scoring with either the Likert or PIRADS v2 scoring systems(1,2,7). ES, as a quantitative score, offers to overcome this limitation. ES is, by definition, little sensitive to volume segmentation.

Performances of the Entropy Score were comparable to Visual Scoring with AUC of 0.89 and 0.91 (p=0.75). With an optimized threshold, ES significantly showed higher proportion (54%) of positive lesion for SPCa than VS (24%), without missing SPCa. Depending of the PIRADS v2 score, rate of false positive generated by VS has been reported ranging from 17% to 88% in a population undergoing targeted biopsy. The positive association of ES from MRI lesions to SPCa offers the possibility to select the MRI generated lesions to target with biopsy.

Automated texture analysis attempts to replicate a deep learning process empowered by computer calculation capabilities, but are highly dependent of quality of input data and their clinical relevance(10). The current workflow leading to computation of the ES can be fully automated.

As a potential tool, longitudinal analysis of ES in an MRI guided active surveillance population might be of interest to detect progression of cancer leading to change in management(51). Even E from other imaging modality, like Positron Emission Tomography scan, can possibly be added and extend performances.

Our choice of a biopsy population permits to avoid the bias of selection of patient undergoing radical prostatectomy, with inclusion of patient potentially eligible for active surveillance and without the diagnosis of PCa. The reference standard, providing ground truth, meets the criteria of standard of care by sampling systematically the whole gland with addition of targeting MRI lesion. In a recent meta-analysis, no difference was shown between cognitive and fusion biopsy for detection rate(52).

Limitations

Regardless the high number of generated imaging volumes, the cohort is relatively small even if some statistical differences have been found either for quantitative values or ES. This early proof of concept study was powered to detect a significant difference in quantitative parameters, not specifically ES. Technically, mpMRI did not include high b value sequences and was acquired at 1.5T. Some refinements might increase the accuracy of ES , for example using 3T scanner to increase the signal to noise ratio. ES for selection of targets

was tested for lesion generated by VS, standard test for reporting mpMRI. A fully automated workflow to compute ES can be implemented and would currently rely on generation of MRI lesion with human assistance, by its segmentation on a single sequence, awaiting validation of automation of this step by the means of artificial neural networks.

Also, the computed threshold in this experiment depends of the chosen conservative definition of clinically significant prostate cancer, including Gleason 6. Variation of this debated definition will probably impact performances(53).

Research implications

This early proof concept study has multiple research implications. Validation of imaging biomarkers has to follow an established process (36). Mainly, association of ES, heterogeneity at imaging, has to be confirmed as meaningful signature of significant prostate cancer based on biological features and ultimately patient outcomes (54). Testing this score in a larger, prospectively acquired biopsy naïve population with a robust reference test is a mandatory. We plan to confront those preliminary findings for detection in the PROMIS trial cohort that fulfils those criteria in a retrospective analysis(1). This is a compulsory step before to undertake a prospective validation study in a new population. For longitudinal analysis, quantitative assessment of progression using ES of MRI lesion in a surveillance setting would first need a longitudinal descriptive study linked to clinical outcomes. This would require a robust histologic gold standard where an MRI lesion is systematically resampled at different time points. A deep understanding of the heterogeneity of prostate tissue, either benign or cancerous, is needed at both histology and biomolecular level of analysis. There is a need of high quality data to input in texture analysis platforms (10). Other and more complex texture analysis features might outperform the entropy and need to be tested (32).

Conclusion

We report results of a proof of concept study which shows association between a quantitative texture-based Entropy Score, linked to heterogeneity across MRI sequences, and SPCa. For similar overall performances to Visual Scoring, ES presented higher proportion of positive lesions for SPCa without compromising NPV and Se. This permits to envisage selection of MRI generated targets for biopsy. Multiple steps are needed to validate those initial findings as an imaging biomarker.

Conflicts of interest

None declared

REFERENCES :

1. Ahmed HU, El-Shater Bosaily A Brown LC, Gabe R, Kaplan R, Parmar MK, et al. Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study. Lancet. 2017 Feb 25;389(10071):815–22.

2. Kasivisvanathan V, Rannikko AS, Borghi M, Panebianco V, Mynderse LA, Vaarala MH, et al. MRI-Targeted or Standard Biopsy for Prostate-Cancer Diagnosis. New England Journal of Medicine [Internet]. 2018 Mar 18 [cited 2018 Mar 19]; Available from: http://www.nejm.org/doi/10.1056/NEJMoa1801993

3. Moldovan PC, Van den Broeck T, Sylvester R, Marconi L, Bellmunt J, van den Bergh RCN, et al. What Is the Negative Predictive Value of Multiparametric Magnetic Resonance Imaging in Excluding Prostate Cancer at Biopsy? A Systematic Review and Meta-analysis from the European Association of Urology Prostate Cancer Guidelines Panel. European Urology. 2017 Aug;72(2):250–66.

4. Martin RM, Donovan JL, Turner EL, Metcalfe C, Young GJ, Walsh EI, et al. Effect of a Low-Intensity PSA-Based Screening Intervention on Prostate Cancer Mortality: The CAP Randomized Clinical Trial. JAMA. 2018 Mar 6;319(9):883.

5. Andriole GL, Crawford ED, Grubb RL, Buys SS, Chia D, Church TR, et al. Prostate cancer screening in the randomized Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial: mortality results after 13 years of follow-up. J Natl Cancer Inst. 2012 Jan 18;104(2):125–32.

6. Schröder FH, Hugosson J, Roobol MJ, Tammela TLJ, Zappa M, Nelen V, et al. Screening and prostate cancer mortality: results of the European Randomised Study of Screening for Prostate Cancer

(ERSPC) at 13 years of follow-up. The Lancet. 2014 Dec 6;384(9959):2027–35.

7. Rosenkrantz AB, Ginocchio LA, Cornfeld D, Froemming AT, Gupta RT, Turkbey B, et al. Interobserver Reproducibility of the PI-RADS Version 2 Lexicon: A Multicenter Study of Six Experienced Prostate Radiologists. Radiology. 2016 Apr 1;280(3):793–804.

8. Azahaf M, Haberley M, Betrouni N, Ernst O, Behal H, Duhamel A, et al. Impact of arterial input function selection on the accuracy of dynamic contrast-enhanced MRI quantitative analysis for the diagnosis of clinically significant prostate cancer: Impact of AIF on K^{trans} in PCa. Journal of Magnetic Resonance Imaging. 2016 Mar;43(3):737–49.

9. Drozdzal M, Chartrand G, Vorontsov E, Shakeri M, Di Jorio L, Tang A, et al. Learning normalized inputs for iterative estimation in medical image segmentation. Medical Image Analysis. 2018 Feb 1;44:1–13.

10. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Medical Image Analysis. 2017 Dec 1;42:60–88.

11. Depeursinge A, Foncubierta-Rodriguez A, Van De Ville D, Müller H. Three-dimensional solid texture analysis in biomedical imaging: Review and opportunities. Medical Image Analysis. 2014 Jan;18(1):176–96.

12. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. Radiology. 2015 Nov 18;278(2):563–77.

13. Andor N, Graham TA, Jansen M, Xia LC, Aktipis CA, Petritsch C, et al. Pan-cancer analysis of the extent and consequences of intratumor heterogeneity. Nat Med. 2016 Jan;22(1):105–13.

14. Epstein JI, Egevad L, Amin MB, Delahunt BM, Srigley JRM, Humphrey PA, et al. The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma: Definition of Grading Patterns and Proposal for a New Grading System. Journal of Surgical Pathology. 2016 Feb;40(2):244–52.

15. Cyll K, Ersvær E, Vlatkovic L, Pradhan M, Kildal W, Kjær MA, et al. Tumour heterogeneity poses a significant challenge to cancer biomarker research. Br J Cancer. 2017 Jul 25;117(3):367–75.

16. Wei L, Wang J, Lampert E, Schlanger S, DePriest AD, Hu Q, et al. Intratumoral and Intertumoral Genomic Heterogeneity of Multifocal Localized Prostate Cancer Impacts Molecular Classifications and Genomic Prognosticators. European Urology. 2017 Feb 1;71(2):183–92.

17. Abeshouse A, Ahn J, Akbani R, Ally A, Amin S, Andry CD, et al. The Molecular Taxonomy of Primary Prostate Cancer. Cell. 2015 Nov 5;163(4):1011–25.

18. Castellano G, Bonilha L, Li LM, Cendes F. Texture analysis of medical images. Clinical Radiology. 2004 Dec;59(12):1061–9.

19. Shannon C. A Mathematical Theory of Communication. 1948 Oct;27:379–423.

20. Kim J-H, Ko ES, Lim Y, Lee KS, Han B-K, Ko EY, et al. Breast Cancer Heterogeneity: MR Imaging Texture Analysis and Survival Outcomes. Radiology. 2016 Oct 4;282(3):665–75.

21. Wibmer A, Hricak H, Gondo T, Matsumoto K, Veeraraghavan H, Fehr D, et al. Haralick

texture analysis of prostate MRI: utility for differentiating non-cancerous prostate from prostate cancer and differentiating prostate cancers with different Gleason scores. European Radiology. 2015 Oct;25(10):2840–50.

22. Rosenkrantz AB, Triolo MJ, Melamed J, Rusinek H, Taneja SS, Deng F-M. Whole-lesion apparent diffusion coefficient metrics as a marker of percentage Gleason 4 component within Gleason 7 prostate cancer at radical prostatectomy. J Magn Reson Imaging. 2014 Feb 25;

23. Sidhu HS, Benigno S, Ganeshan B, Dikaios N, Johnston EW, Allen C, et al. "Textural analysis of multiparametric MRI detects transition zone prostate cancer". Eur Radiol. 2016 Sep 12;1–11.

24. Rosenkrantz AB, Ream JM, Nolan P, Rusinek H, Deng F-M, Taneja SS. Prostate Cancer: Utility of Whole-Lesion Apparent Diffusion Coefficient Metrics for Prediction of Biochemical Recurrence After Radical Prostatectomy. American Journal of Roentgenology. 2015 Dec;205(6):1208–14.

25. Rosenkrantz AB, Mendrinos S, Babb JS, Taneja SS. Prostate cancer foci detected on multiparametric magnetic resonance imaging are histologically distinct from those not detected. J Urol. 2012 Jun;187(6):2032–8.

26. Barentsz JO, Richenberg J, Clements R, Choyke P, Verma S, Villeirs G, et al. ESUR prostate MR guidelines 2012. Eur Radiol. 2012 Apr;22(4):746–57.

27. Rosenkrantz AB, Kim S, Lim RP, Hindman N, Deng F-M, Babb JS, et al. Prostate Cancer Localization Using Multiparametric MR Imaging: Comparison of Prostate Imaging Reporting and Data System (PI-RADS) and Likert Scales. Radiology. 2013 Jun 20;

28. Barentsz J, Villers A, Schouten M. Reply to Letter to the Editor re: ESUR prostate MR guidelines. European Radiology. 2013 Jun 11;23(8):2322–3.

29. Haffner J, Lemaitre L, Puech P, Haber G-P, Leroy X, Jones JS, et al. Role of magnetic resonance imaging before initial biopsy: comparison of magnetic resonance imaging-targeted and systematic biopsy for significant prostate cancer detection: *ROLE OF MRI BEFORE INITIAL BIOPSY*. BJU International. 2011 Oct;108(8b):E171–8.

30. Epstein JI, Walsh PC, Carmichael M, Brendler CB. Pathologic and Clinical Findings to Predict Tumor Extent of Nonpalpable (Stage T1 c) Prostate Cancer. JAMA: The Journal of the American Medical Association. 1994 Feb 2;271(5):368–74.

31. Depeursinge A, Foncubierta-Rodriguez A, Van De Ville D, Müller H. Three-dimensional solid texture analysis in biomedical imaging: Review and opportunities. Medical Image Analysis. 2014 Jan;18(1):176–96.

32. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nature Reviews Clinical Oncology. 2017 Oct 4;14(12):749–62.

33. Milletari F, Navab N, Ahmadi S-A. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. In IEEE; 2016 [cited 2018 Jan 24]. p. 565–71. Available from: http://ieeexplore.ieee.org/document/7785132/

34. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF, editors. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015 [Internet]. Cham: Springer International Publishing; 2015 [cited 2018 Jan 24]. p. 234–41. Available from: http://link.springer.com/10.1007/978-3-319-24574-4_28

35. O'Connor JPB, Aboagye EO, Adams JE, Aerts HJWL, Barrington SF, Beer AJ, et al. Imaging biomarker roadmap for cancer studies. Nature Reviews Clinical Oncology. 2016 Oct 11;14(3):169–86.

36. van Niekerk CG, van der Laak JAWM, Börger ME, Huisman H-J, Witjes JA, Barentsz JO, et al. Computerized whole slide quantification shows increased microvascular density in pT2 prostate cancer as compared to normal prostate tissue. Prostate. 2009 Jan;69(1):62–9.

37. van Niekerk CG, van der Laak JAWM, Hambrock T, Huisman H-J, Witjes JA, Barentsz JO, et al. Correlation between dynamic contrast-enhanced MRI and quantitative histopathologic microvascular parameters in organ-confined prostate cancer. European Radiology. 2014 Oct;24(10):2597–605.

38. Truong M, Hollenberg G, Weinberg E, Messing EM, Miyamoto H, Frye TP. Impact of Gleason Subtype on Prostate Cancer Detection Using Multiparametric Magnetic Resonance Imaging: Correlation with Final Histopathology. The Journal of Urology. 2017 Aug 1;198(2):316–21.

39. Hectors SJ, Wagner M, Bane O, Besa C, Lewis S, Remark R, et al. Quantification of hepatocellular carcinoma heterogeneity with multiparametric magnetic resonance imaging. Scientific Reports [Internet]. 2017 Dec [cited 2017 Aug 8];7(1). Available from: http://www.nature.com/articles/s41598-017-02706-z

40. Zhang B, Tian J, Dong D, Gu D, Dong Y, Zhang L, et al. Radiomics Features of Multiparametric MRI as Novel Prognostic Factors in Advanced Nasopharyngeal Carcinoma. Clinical Cancer Research. 2017 Aug 1;23(15):4259–69.

41. Choy B, Pearce SM, Anderson BB, Shalhav AL, Zagaja G, Eggener SE, et al. Prognostic Significance of Percentage and Architectural Types of Contemporary Gleason Pattern 4 Prostate Cancer in Radical Prostatectomy: The American Journal of Surgical Pathology. 2016 Oct;40(10):1400–6.

42. Oto A, Yang C, Kayhan A, Tretiakova M, Antic T, Schmid-Tannwald C, et al. Diffusion-Weighted and Dynamic Contrast-Enhanced MRI of Prostate Cancer: Correlation of Quantitative MR Parameters With Gleason Score and Tumor Angiogenesis. American Journal of Roentgenology. 2011 Dec 1;197(6):1382–90.

43. Verma S, Rajesh A, Morales H, Lemen L, Bills G, Delworth M, et al. Assessment of Aggressiveness of Prostate Cancer: Correlation of Apparent Diffusion Coefficient With Histologic Grade After Radical Prostatectomy. American Journal of Roentgenology. 2011 Jan 21;196(2):374–81.

44. Vignati A, Mazzetti S, Giannini V, Russo F, Bollito E, Porpiglia F, et al. Texture features on T2-weighted magnetic resonance imaging: new potential biomarkers for prostate cancer aggressiveness. Phys Med Biol. 2015 Apr 7;60(7):2685–701.

45. Ginsburg SB, Algohary A, Pahwa S, Gulani V, Ponsky L, Aronen HJ, et al. Radiomic features

for prostate cancer detection on MRI differ between the transition and peripheral zones: Preliminary findings from a multi-institutional study: Radiomic Features for Prostate Cancer Detection on MRI. Journal of Magnetic Resonance Imaging. 2017 Jul;46(1):184–93.

46. Rosenkrantz AB, Meng X, Ream JM, Babb JS, Deng F-M, Rusinek H, et al. Likert score 3 prostate lesions: Association between whole-lesion ADC metrics and pathologic findings at MRI/ultrasound fusion targeted biopsy: Whole-Lesion ADC Metrics in Likert 3 Prostate Lesions. Journal of Magnetic Resonance Imaging. 2016 Feb;43(2):325–32.

47. Dickinson L, Ahmed HU, Allen C, Barentsz JO, Carey B, Futterer JJ, et al. Magnetic resonance imaging for the detection, localisation, and characterisation of prostate cancer: recommendations from a European consensus meeting. Eur Urol. 2011 Apr;59(4):477–94.

48. Weinreb JC, Barentsz JO, Choyke PL, Cornud F, Haider MA, Macura KJ, et al. PI-RADS Prostate Imaging – Reporting and Data System: 2015, Version 2. European Urology. 2016 Jan;69(1):16–40.

49. Fütterer JJ, Briganti A, De Visschere P, Emberton M, Giannarini G, Kirkham A, et al. Can Clinically Significant Prostate Cancer Be Detected with Multiparametric Magnetic Resonance Imaging? A Systematic Review of the Literature. European Urology. 2015 Dec;68(6):1045–53.

50. Moore CM, Giganti F, Albertsen P, Allen C, Bangma C, Briganti A, et al. Reporting Magnetic Resonance Imaging in Men on Active Surveillance for Prostate Cancer: The PRECISE Recommendations-A Report of a European School of Oncology Task Force. Eur Urol. 2017 Apr;71(4):648–55.

51. Wegelin O, van Melick HHE, Hooft L, Bosch JLHR, Reitsma HB, Barentsz JO, et al. Comparing Three Different Techniques for Magnetic Resonance Imaging-targeted Prostate Biopsies: A Systematic Review of In-bore versus Magnetic Resonance Imaging-transrectal Ultrasound fusion versus Cognitive Registration. Is There a Preferred Technique? European Urology. 2017 Apr;71(4):517–31.

52. Netto GJ. Tumor Volume Threshold of Insignificant Prostate Cancer--Was Dr. Stamey Right All Along? The Journal of Urology. 2011 Jan;185(1):10–1.

53. D'Amico AV. Active Surveillance Versus Treatment of Prostate Cancer: Should Metastasis Be the Primary End Point? J Clin Oncol. 2017 May 20;35(15):1638–40.

Tables

Table 1. Included Patients characteristics

Number patients	20
Median age (years, range)	65 (55-74)
Median PSA (ng/ml, range)	8.22 (4.54-52.72)
Number of patients with positive biopsy	12
Number of patients under active surveillance	2
Number of patients biopsy naive	14
Number of patients with previous negative TRUS biopsy	4
Number of high risk patients	2
Number of patient mpMRI naive	20
Maximum Gleason score per patient	
Gleason 6 (3+3)	8
Gleason 7 (3+4)	3
Gleason 8 (4+4)	1
Median number of target per patient (range)	1.5 (0-3)
Median maximum target score per patient (range)	3 (3-5)
Number of positive targets for any cancer (n significant cancer)	
score 3/5	2/19 (0)
score 4/5	2/5 (2)
score 5/5	3/3 (3)
Mean Maximum core length (range)	5.5 (1-17)

Table 2. Values of Entropy Score for different VOIS:

	Entropy Score			
Volume of interest (n)	mean	median	standard deviation	CI 95%
Prostate (20)	18.31	18.78	1.17	0.51
whole Peripheral Zone (20)	17.96	18.56	1.30	0.57
whole Transition Zone (20)	17.96	18.20	1.26	0.55
Normal Peripheral Zone (19)***	**15.58**	15.47	3.84	1.72
BPH nodule (18)***	**16.26**	16.38	5.30	2.45
Target (28)	16.31	16.31	1.22	0.45
Positive Target (5)*	**17.21**	16.87	1.00	0.42
Negative Target (23)*	**16.11**	15.88	1.19	0.50
Cancer (12)	16.93	16.81	1.59	0.90
Significant Cancer (8)**	**17.73**	17.40	1.23	0.85
Non-significant cancer (4)**	**15.33**	15.12	0.78	0.76

*p=0.035; **p=0.0025; ***p=0.04

Table 3. Comparison of performances of the Entropy Score and Multi logistic regression models:

	A analysis All VOIs		B Analysis-MRI Lesions	
	Entropy Score	Regression Model	Entropy Score	Regression Model
Sensitivity % (95% CI)	100 (63.0 - 100)	87.5 (47.3 - 99.6)	100 (63.0 - 100)	87.5 (47.3 - 99.6)
Specificity % (95% CI)	69.3 (56.3 - 80.4)	88.71 (78.1 - 95.3)	72 (50.6 - 87.9)	84.3 (63.9 - 95.4)
PPV % (95% CI)	29.6 (13.7 - 50.1)	50 (95% CI 23.0 - 76.)	53.3 (26.5 - 78.7)	63.6 (30.7 - 89.0)
NPV % (95% CI)	100 (91.7 - 100)	98.2 (90.4 - 99.9)	100 (81.4 - 100)	95.4 (77.1 - 99.8)
Diagnostic Accuracy % (95% CI)	72.8 (60.9 - 82.8)	88.5 (78.7 - 94.9)	78.7 (61.0 - 91.0)	84.8 (68.1 - 94.8)
Youden index (95% CI)	0.69 (0.19 - 0.8)	0.76 (0.25 - 0.95)	0.72 (0.14 - 0.88)	0.72 (0.11 - 0.95)
AUC (95% CI)	0.88 (0.76, 0.97)	0.93 (0.84, 0.99)	0.89 (0.76, 0.99)	0.93 (0.82, 1.00)
p value for AUC comparison	0.16		0.32	

PPV: positive predictive value; NPV: negative predictive value; CI: confidence interval; AUC: area under the curve; VOI: volume of interest.

Table 4: Contingency table for lesion detection by VS and ES

	Visual Scoring	Entropy Score	
	3,4 and 5	<16.61Nat	>16.61Nat
Positive for SPCa	8	0	8
Negative for SPCa	25	18	7
Number of lesions	33	18	15
Detection rate	**0.24***	0	**0.53***

***significant difference p<0.0001

Figures Legends :

Figure 1. Study profile and different levels of statistical analysis with matching investigated population

Figure 2 ROC curves. (**A)** for detection of significant cancer using entropy score and each of its component within the MRI lesions and (**B)** ROC curves for ES (AUC 0.89) and VS (0.91) are plotted for B analysis(p=0.75)

Figure 3 shows two examples of high and low entropy score of cancerous lesions of Gleason (3+3)in A and (3+4) in B with Entropy Score of 14.71 NAT and 18.75 NAT respectively.