diagnostics

MDPI

Article

Advanced Imaging Analysis in Prostate MRI: Building a Radiomic Signature to Predict Tumor Aggressiveness

Anna Damascelli [1,†], Francesca Gallivanone [2,†], Giulia Cristel [1], Claudia Cava [2], Matteo Interlenghi [2], Antonio Esposito [1,3], Giorgio Brembilla [1], Alberto Briganti [3,4], Francesco Montorsi [3,4], Isabella Castiglioni [5,*] and Francesco De Cobelli [1,3]

1 Department of Radiology, IRCCS San Raffaele Scientific Institute, 20132 Milan, Italy; damascelli.anna@hsr.it (A.D.); cristel.giulia@hsr.it (G.C.); esposito.antonio@hsr.it (A.E.); brembilla.giorgio@hsr.it (G.B.); decobelli.francesco@hsr.it (F.D.C.)
2 Institute of Molecular Bioimaging and Physiology, National Research Council (IBFM-CNR), 20090 Segrate, Italy; francesca.gallivanone@ibfm.cnr.it (F.G.); claudia.cava@ibfm.cnr.it (C.C.); matteo.interlenghi@ibfm.cnr.it (M.I.)
3 Vita-Salute San Raffaele University, 20132 Milan, Italy; briganti.alberto@hsr.it (A.B.); montorsi.francesco@hsr.it (F.M.)
4 Department of Urology, IRCCS San Raffaele Scientific Institute, 20132 Milan, Italy
5 Department of Physics "G. Occhialini", University of Milano, 20126 Bicocca, Italy
* Correspondence: isabella.castiglioni@unimib.it; Tel.: +39-022-171-7511
† Authors equally contributed and share first authorship.

Abstract: Radiomics allows the extraction quantitative features from imaging, as imaging biomarkers of disease. The objective of this exploratory study is to implement a reproducible radiomic-pipeline for the extraction of a magnetic resonance imaging (MRI) signature for prostate cancer (PCa) aggressiveness. One hundred and two consecutive patients performing preoperative prostate multiparametric magnetic resonance imaging (mpMRI) and radical prostatectomy were enrolled. Multiparametric images, including T2-weighted (T2w), diffusion-weighted and dynamic contrast-enhanced images, were acquired at 1.5 T. Ninety-three imaging features (Ifs) were extracted from segmentation of index lesion. Ifs were ranked based on a stability rank and redundant Ifs were excluded. Using unsupervised hierarchical clustering, patients were grouped on the basis of similar radiomic patterns, whose association with Gleason Grade Group (GGG), extracapsular extension (ECE), and nodal involvement (pN) was tested. Signatures composed by IFs from T2w-images and Apparent Diffusion Coefficient (ADC) maps were tested for the prediction of GGG, ECE, and pN. T2w radiomic pattern was associated with pN, ECE, and GGG (p = 0.027, 0.05, 0.03) and ADC radiomic pattern was associated with GGG (p = 0.004). The best performance was reached by the signature combing IFs from multiparametric images (0.88, 0.89, and 0.84 accuracy for GGG, pN, and ECE). A reliable multiparametric MRI radiomic signature was extracted, potentially able to predict PCa aggressiveness, to be further validated on an independent sample.

Keywords: magnetic resonance imaging; prostate cancer; prostate cancer aggressiveness; radiomics

check for updates

Citation: Damascelli, A.; Gallivanone, F.; Cristel, G.; Cava, C.; Interlenghi, M.; Esposito, A.; Brembilla, G.; Briganti, A.; Montorsi, F.; Castiglioni, I.; et al. Advanced Imaging Analysis in Prostate MRI: Building a Radiomic Signature to Predict Tumor Aggressiveness. *Diagnostics* **2021**, *11*, 594. https://doi.org/10.3390/diagnostics11040594

Received: 19 February 2021
Accepted: 24 March 2021
Published: 26 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Prostate cancer (PCa) is a heterogeneous disease characterized by a wide spectrum of clinical presentations and possible outcomes [1]. In clinical practice, the main challenge is to identify the best balance between limiting overtreatment and reducing mortality.

In the preoperative setting, PCa aggressiveness is commonly established using Gleason grade group (GGG) derived from prostate biopsy specimens that, in 30–50% of cases, does not represent the true GGG of the tumor [2–4]. The identification of a non-invasive and accurate tool to predict tumor aggressiveness, taking into account the biological heterogeneity present in whole tumor volume, is an important unmet need.

In the last decade, Radiomics emerged as a post-processing imaging analysis that allows the extraction of large number of quantitative features that cannot be studied solely by visual assessment [5,6]. The role of radiomic in oncological setting is to identify reliable biomarkers of cancer aggressiveness [7,8]. In recent years, several studies explored a radiomic approach applied to multiparametric magnetic resonance imaging (mpMRI) for detection and characterization of PCa [8,9]. Despite the recommendations of international working groups to develop a standardized pipeline to extract image biomarkers, they are characterized by a great variability in methodological approach and results. Furthermore, most of these studies focused the analysis on the prediction of GGG solely, while the association of radiomic features with other important prognostic parameters, such as extracapsular extension (ECE) and nodal involvement, remains quite unexplored.

This study aimed to implement a reproducible radiomic-pipeline for the extraction of a radiomic signature from mpMRI T2w images and Apparent Diffusion Coefficient (ADC) maps able to predict Pca aggressiveness, in terms of GGG, ECE, and nodal stage (pN).

2. Materials and Methods

This is a retrospective study approved by our Institutional Review Board; written informed consent was obtained from all patients.

2.1. Study Population

The study cohort consisted of 102 men with biopsy proven PCa who underwent both preoperative mpMRI of the prostate and radical prostatectomy at a single tertiary care referral center, between January 2016 and March 2019.

Pathological data of surgical specimen included Gleason grade group (GGG), extracapsular extension (ECE), and nodal stage (pN).

Prostate specimens were processed according to the Stanford protocol [10] and analyzed by a dedicated uro-pathologist with 20 years of experience.

2.2. MRI Protocol

MRI examinations were performed using a single 1.5 T scanner (Achieva dStream, Philips Medical Systems, Best, The Netherlands) with a balloon-covered expandable endorectal coil (BPX-15™, Bayer Medical Care, Indianola, PA, USA) paired with a phased array 32-channel surface coil. Gastrointestinal peristalsis was suppressed by intramuscular administration of 20 mg of scopolamine-butylbromide. Detailed mpMRI imaging protocol is reported in Table 1. Data from DCE-MRI (intravenous bolus injection of 0.1 mmol/kg of gadobutrol) were acquired for clinical needs but not considered for quantitative analysis.

Table 1. Imaging protocol details.

Parameter	T2 TSE * Axial	T2 TSE * Sagittal	T2 TSE * Coronal	DWI ** (b: 0, 800, 1600)	DCE ***
TR (ms)	4824	4370	2991	4376	3.7
TE (ms)	120	120	120	80	1.83
FOV° (mm)	180 × 180	180 × 180	180 × 180	180 × 180	180 × 180
Matrix Thickness (mm)	3	3	3	3	3
Gap (mm)	0.3	0.3	0.3	0.3	0
Flip angle (°)	90	90	90	90	8/5, 8, 12, 15
Acquisition time	4 min 6 s	3 min 25 s	2 min 8 s	5 min 19 s	3 min 20 s

* TSE: Turbo spin echo imaging; ** DWI: Diffusion-weighted imaging; *** DCE: Dynamic contrast-enhanced imaging; ° FOV: Field of view.

2.3. MRI Qualitative Analysis

Two dedicated radiologists (5 and 4 years of experience in prostate mpMRI) independently reviewed the images and identified the index lesion based on PIRADS v. 2.0 classification. Images were than compared with prostatectomy reports to check the concordance between imaging and pathology findings. In all cases, the index lesion at MRI matched the

lesion with highest GGG at pathology. In case of multiple lesions at MRI, with same GGG at pathology, the largest one in the images was selected for the analysis.

2.4. MRI Quantitative Analysis

2.4.1. MRI Images Post-Processing

T2w images were pre-processed to account for the impact of bias field and intensity non-standardness on T2w images [11], by N4 normalization [12] and linear scaling and shifting within 0–600 in 3D Slicer v. 4.10.2 [13].

Diffusion weighted images were processed using monoexponential model with *b*-values = 50 s/mm^2 and *b*-values = 800 s/mm^2, to obtain ADC maps.

2.4.2. Tumor Volume Segmentation

Index lesions were segmented both on T2w images and ADC maps using a semi-automatic approach based on region growing [14], as implemented in 3D Slicer v. 4.10.2 [13], and the 3D lesion volume (volume of interest (VOI)) was obtained. In order to avoid registration errors as well as motion artifacts, tumors were segmented separately on T2w images and ADC maps (see Figure 1).

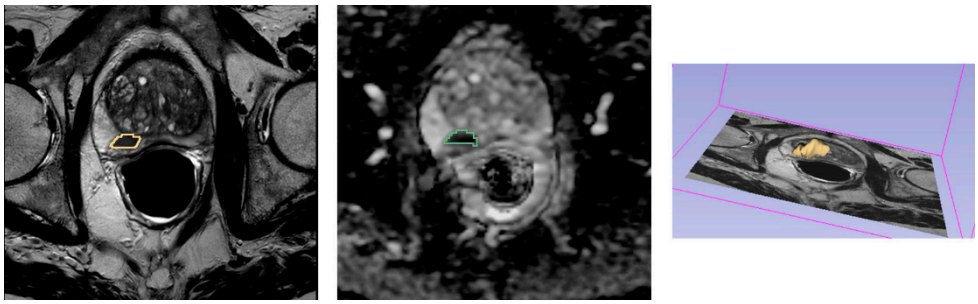

Figure 1. Example of an index lesion segmentation on T2w image (**left**) and ADC map (**center**) and the resulted 3D segmented volume of interest (VOI) (**right**).

As a preliminary step, to evaluate the stability of the quantitative imaging features with respect to segmentation, two radiologists independently performed tumor delineation. For subsequent analysis in a separate session, the two readers re-segmented in consensus the images of the entire patient sample.

Only patients with segmented VOI greater 0.7 cc in ADC were included in the subsequent analysis. Dimensional cut-off on images were empirically evaluated, considering that a lesion with less than 4 voxels in each direction would not make for meaningful calculation of different textural parameters [6].

2.4.3. MRI Image Quantification

Feature extraction was performed by using Radiomics module of 3D slicer v. 4.10.2 [15]. Segmented VOIs were resampled to isotropic voxel spacing, using an upsampling scheme based image slice thickness. Image re-segmentation was performed with Collewet normalization [16] and VOI intensities were discretized to a fixed number of 64 bins.

A total number of 93 imaging features (IFs) was extracted from both T2w images (T2w-IFs) and ADC map (ADC-IFs), without applying filters, from five different classes (14 morphological features (M), 18 first-order statistical features from intensity histogram (FOS), including mean ADC value, and 61 textural features from analysis of Gray Level Co-occurrence Matrix (GLCM), Gray Level Size Zone Matrix (GLSZM), Gray Level Run Length Matrix (GLRLM), and Neighborhood Grey Tone Difference Matrix (NGTDM). Details on radiomic pipeline for feature extraction can be found in Table S1 of Supplementary Materials.

To evaluate the impact of semi-automatic segmentation on image quantification, the nonparametric repeated measurement Friedman test was used to calculate a stability rank per feature from independent segmentations performed by two readers [17,18]. Further-

more, feature selection was carried out on the extracted IFs for each image modality in order to rule out redundant features, by analyzing the covariance matrix. Clusters of highly correlated features ($r > 0.8$) were reduced to a single representative imaging feature, using the one presenting the highest inter-subject's range [6], calculated by Coefficient of Variation [17,18]. For subsequent analysis, only non-redundant IFs, on the basis of the stability rank for delineation inaccuracies by different operators, were retained.

2.4.4. Association of Radiomic Phenotype and Clinical Data

Unsupervised hierarchical clustering analysis was performed to group patients on the basis of similar radiomic patterns. Each patient was considered a statistical unit (sample) characterized by different measurable properties (IFs values of patient's index lesion). The clustering procedure grouped the patients showing a closer radiomic pattern, as defined by a distance criterion (Euclidean distance).

Before the clustering analysis, IFs presenting extreme values were log2 transformed, replacing zero values with the smallest positive value of the radiomic phenotype, over all patient samples, before transformation. IFs were standardized to zero mean and a unit standard deviation. Heatmaps were automatically generated by using ComplexHeatmap in R-package [19].

Fisher's exact test as implemented was used to test statistical significance of the association between identified clusters and patient histological characteristics.

2.4.5. Building Diagnostic Radiomic Signature

Different signatures composed by IFs from T2w images and ADC maps independently and by IFs from the two modality jointly were built and tested for prediction of high ($GGG \geq 4 + 3$) vs. low ($GGG < 4 + 3$) GGG, presence vs. absence of ECE, and pN status (pN0 vs. $pN \geq 1$). For each image modality we selected only non-redundant IFs within the first 10 stable Ifs in the stability rank (S_{TOP}). ADC_{mean} was evaluated both separately and in combination with Ifs signatures, to test the effective potential advantage of textural features compared to standard value ADC_{mean}. A Support Vector Machine model was implemented using a R-package [20,21] with kernel radial in order to evaluate the combination of features that achieved the best performances in the prediction model. Classification performances were evaluated in terms of accuracy, sensitivity, and specificity as metrics. Since an imbalance exists between the different classes to be considered in the classification, we employed data resampling, with an undersampling method [22,23]. We randomly resampled the cohort, removing samples from the majority class in order to obtain 10 random subsets with the same number of samples in the two classes. Then we performed the classification in each of these subsets, so the classification was performed 10 times. The overall mean value of accuracy, specificity, and sensitivity was obtained as means over the 10 repetitions.

Since the biology upon peripheral and transition zone lesions could be different across characteristics, the classification was performed also in the subset of PZ lesions, in order to evaluate the performance of signatures in this subgroup. Considering the low number of TZ lesions, no classification was made on these samples.

3. Results

From the initial population of 102 patients, 40 of them were excluded according to the imaging volumetric criteria (VOI < 0.7 cc in ADC map). The final population included for the analysis consisted of 62 patients (age range 44–79 years) with 62 index lesions.

Lesions characteristics at MRI and pathological analysis are reported in Table 2.

Table 2. Index lesions imaging and pathological characteristics.

		# Patients	Frequency
Index lesion location	PZ *	43	69%
	TZ **	16	26%
	Both	3	5%
PI-RADS ***	3	3	5%
	4	30	48%
	5	29	47%
Gleason Score	7 (3 + 4)	18	29%
	7 (4 + 3)	18	29%
	8	5	8%
	9	21	34%
ECE °	Yes	38	61%
	No	24	39%
pN	pN0 °°	39	63%
	pN1 ≥ 1 °°°	13	21%
	pNx #	10	16%

* PZ: Peripheral Zone; ** TZ: Transition Zone; *** PI-RADS: Prostate Imaging-Reporting and Data System; ° ECE: Extracapsular extension; °° pN0: absence of nodal metastases at pathologic examination, °°° pN1: presence of nodal metastases at pathologic examination. # pNx: nodes status not assessable.

On a subset of patients, the index lesion was segmented by two operators, and a stability rank with respect to segmentation was defined for all the 93 IFs extracted from both T2w images and ADC maps. Forty-four out of ninety-three T2w-IFs and 61/93 ADC-IFs resulted in being stable with respect to semi-automatic segmentation (p-value of Friedman test > 0.05). Among them, the analysis of covariance matrix allowed to select 17 T2w-IFs and 23 ADC-IFs as not-redundant futures used for further multivariate analysis.

Unsupervised hierarchical clustering revealed two major groups of patients both on the basis of T2w radiomic pattern as well as on the basis of ADC radiomic pattern (Figure 2) or considering IFs from both T2w images and ADC maps (Figure 3).

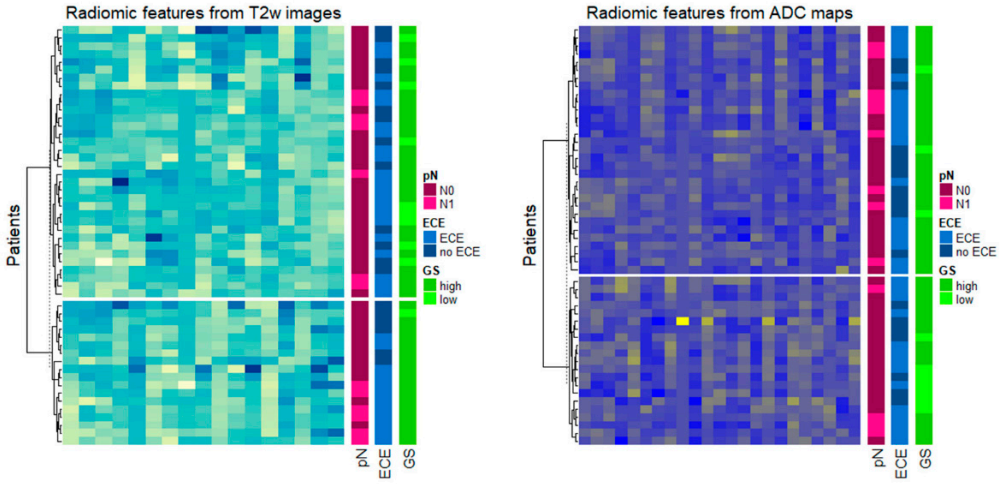

Figure 2. Radiomic patterns obtained on T2w images (**left**) and ADC maps (**right**); the dendrogram at the left side represents the patient' grouping obtained from the clustering procedure; annotation at the right side represent the distribution of Gleason Grade Group (GGG), extracapsular extension (ECE), and nodal stage (pN) status in the groups.

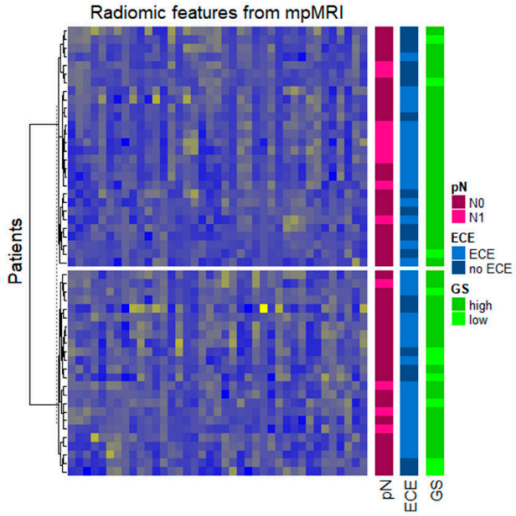

Figure 3. Radiomic patterns obtained considering both T2w images and ADC maps; the dendrogram at the left side represents the patient' grouping obtained from the clustering procedure; annotation at the right side represent the distribution of GGG, ECE, and pN status in the groups.

T2w radiomic pattern was found to be associated with pN, ECE, and GGG (*p*-value: 0.027, 0.05, and 0.03, Fisher's exact test), while the ADC maps' radiomic pattern was found to be associated with GGG (*p*-value: 0.04, Fisher's exact test).

When combining imaging biomarkers from T2w images and ADC maps, the multimodal radiomic patterns of the two patient groups identified by unsupervised hierarchical clustering analysis was found to be associated with GGG (*p*-value: 0.05 Fisher's exact test). To build and test a diagnostic radiomic signature, we focused the analysis on a reduced set of stable and not-redundant imaging features. Table 3 shows different signature models, obtained from T2w images and ADC maps independently and from the two modalities jointly with and without the inclusion of ADC_{mean}, together with models' performance (mean, max, min, and standard deviation of performances obtained by data resampling) in predicting GGG, ECE, and nodal involvement.

Figure 4 shows ROC curves and AUC for classification.

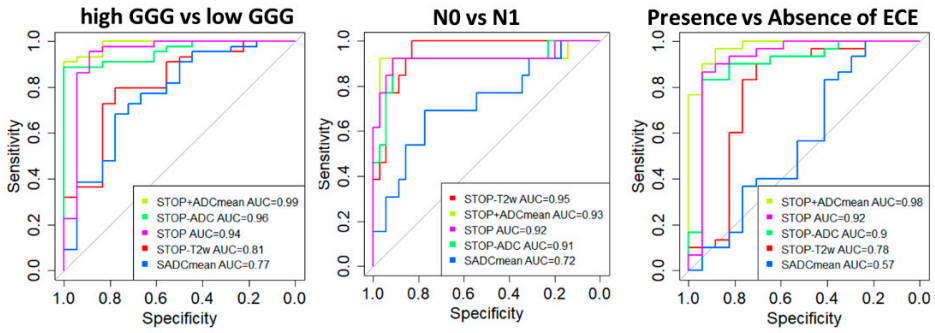

Figure 4. ROC curves and AUC values for classification.

The best performance was reached by the signature combing IFs from T2w images and ADC map ($S_{TOP\text{-}T2w/ADC}$) with an accuracy in the prediction of GGG, nodal involvement, and ECE of 0.88, 0.9m, and 0.85, respectively. The performance was even improved when ADC_{mean} was added to the model (accuracy: 0.90, 0.89, 0.88). Considering each single signature, similar performances were obtained for the dataset of PZ lesions, except for $S_{ADCmean}$, which increases significantly, confirming the predictive role of DWI in PZ lesions (see Table S2 and Figure S1 in Supplementary Materials). Despite this, a slightly higher accuracy has been found for the multimodal signature S_{TOP} (see Table S2 in Supplementary Materials).

Table 3. Signature models and performances. Different signature models obtained from T2w images and ADC maps independently and from the two modalities jointly and the corresponding diagnostic performances. On the left, the specific imaging features (IFs) included in each model are reported. Accuracy, sensitivity, and specificity are reported as Mean and Standard Deviation, while in parenthesis, minimum and maximum over 10 repetitions are reported.

Signature	Image Modality	Feature Group	Features	High GGG vs. Low GGG			N0 vs. N1			Presence vs. Absence of ECE		
				Acc #	Sens ##	Spec ###	Acc #	Sens ##	Spec ###	Acc #	Sens ##	Spec ###
$S_{TOP\text{-}T2w}$ *	T2w images	M	Sphericity	0.75 ± 0.05 (0.67–0.83)	0.73 ± 0.11 (0.56–0.89)	0.76 ± 0.14 (0.44–0.94)	0.84 ± 0.05 (0.77–0.92)	0.79 ± 0.09 (0.62–0.92)	0.88 ± 0.09 (0.69–1)	0.75 ± 0.03 (0.71–0.79)	0.8 ± 0.07 (0.76–0.94)	0.69 ± 0.12 (0.47–0.94)
		GLRLM	SRHGLE									
		GLSZM	SAHGLE									
			LAHGLE									
$S_{TOP\text{-}ADC}$ **	ADC maps	M	Elongation	0.83 ± 0.06 (0.75–0.94)	0.95 ± 0.04 (0.89–1)	0.7 ± 0.13 (0.5–0.89)	0.86 ± 0.05 (0.73–0.92)	0.77 ± 0.1 (0.62–0.92)	0.95 ± 0.09 (0.77–1)	0.81 ± 0.02 (0.79–0.85)	0.85 ± 0.1 (0.65–1)	0.78 ± 0.09 (0.59–0.94)
			Flatness									
		GLCM	Inverse Variance									
			Cluster Shade									
		NGTDM	Busyness									
S_{TOP} ***	T2w images	M	Sphericity	0.88 ± 0.04 (0.81–0.94)	0.94 ± 0.04 (0.89–1)	0.82 ± 0.1 (0.67–0.89)	0.9 ± 0.04 (0.85–0.96)	0.9 ± 0.1 (0.69–1)	0.9 ± 0.1 (0.69–1)	0.85 ± 0.04 (0.79–0.91)	0.93 ± 0.06 (0.88–1)	0.78± 0.09 (0.65–0.94)
		GLRLM	SRHGLE									
		GLSZM	SAHGLE									
			LAHGLE									
	ADC aps	M	Elongation									
			Flatness									
		GLCM	Inverse Variance									
			Cluster Shade									
		NGTDM	Busyness									
$S_{ADCmean}$ °	ADC maps	-	ADC_{mean}	0.71 ± 0.04 (0.64–0.78)	0.64 ± 0.1 (0.44–0.78)	0.78 ± 0.05 (0.72–0.83)	0.67 ± 0.06 (0.54–0.73)	0.51 ± 0.15 (0.23–0.77)	0.82 ± 0.14 (0.62–1)	0.63 ± 0.03 (0.59–0.68)	0.49 ± 0.16 (0.29–0.76)	0.76 ± 0.19 (0.41–1)

Table 3. *Cont.*

Signature	Image Modality	Feature Group	Features	High GGG vs. Low GGG			N0 vs. N1			Presence vs. Absence of ECE		
				Acc [#]	Sens [##]	Spec [###]	Acc [#]	Sens [##]	Spec [###]	Acc [#]	Sens [##]	Spec [###]
$S_{TOP + ADC\ mean}$ [°°]	T2w images	M	Sphericity	0.9 ± 0.04 (0.83–0.94)	0.93 ± 0.06 (0.83–1)	0.88 ± 0.06 (0.78–0.94)	0.89 ± 0.04 (0.81–0.92)	0.82 ± 0.11 (0.62–1)	0.96 ± 0.06 (0.85–1)	0.88 ± 0.03 (0.82–0.91)	0.88 ± 0.07 (0.71–1)	0.89 ± 0.07 (0.76–1)
		GLRLM	SRHGLE									
		GLSZM	SAHGLE									
			LAHGLE									
	ADC maps	M	Elongation									
			Flatness									
		GLCM	Inverse Variance									
			Cluster Shade									
		NGTDM	Busyness									
		-	ADC_{mean}									

* $S_{TOP-T2w}$: S_{TOP} signature extracted from T2w images; ** $S_{TOP-ADC}$: S_{TOP} signature extracted from ADC maps; *** S_{TOP}: S_{TOP} signature extracted from both T2w images and ADC maps; ° $S_{ADCmean}$: signature based on ADC_{mean}; °° $S_{TOP + ADCmean}$: signature obtained by combining S_{TOP} and ADC_{mean}; [#] Acc: accuracy; [##] Sens: sensitivity; [###] Spec: specificity.

4. Discussion

Due to the wide spectrum of clinical presentation and possible outcome of PCa cancer, identifying the optimal treatment option for patients represents the most challenging problem in PCa clinical workup. There is a growing attention in finding reliable tools that could provide a non-invasive assessment of PCa aggressiveness, allowing for a patient-tailored management, ranging from radical-prostatectomy with nodes dissection to active surveillance. Radiomics can analyze a large number of imaging features that cannot be visualized by radiologists, and its potential role in the implementation of tumor detection, characterization, and treatment response has been explored in several oncological settings [24–26]. A large variability of methodological approaches exists among previous studies exploring the application of radiomics to prostate mpMRI. Moreover, the analysis are usually based on single MR quantitative parameters while a multiparametric signature would be preferred, reflecting the multiparametric approach of qualitative clinical reporting [8,9]. From a methodological point of view, our pipeline was designed and standardized in order to identify accurate and reproducible imaging biomarkers addressing the different issues related to the extraction of high-throughput quantitative biomarkers, as underlined by IBSI initiative [27].

The first issue was represented by lesion volume definition. Even if advanced approaches were developed in recent years for PCa segmentation [28], in PCa, as in other oncological diseases, manually contouring is the standard method to define lesion volume [26–28], but, due to its dependence from operator expertise [6], it has been proved to be suboptimal in radiomic studies. In different studies, e.g., [29], lesion volume is defined on a single image slice, thus limiting the fully characterization of volumetric extension of lesions. In order to overcome these limitations, we used a semi-automatic segmentation approach based on region growing on T2w images and ADC maps separately, followed by manual adjustment, as suggested by IBSI.

In order to warrant the reproducibility of IFs included in the radiomic signature, we evaluated a rank of inter-observer stability for IFs extracted from tumor volumes from both T2w images and ADC maps [6,17]. The subsequent analysis was then performed on a set of robust and reproducible IFs. Furthermore, the use of a single MR scanner using a standardized acquisition protocol for all involved patients allowed to avoid issues related to IFs dependency from scanner and acquisition protocols.

Several studies described the association of radiomic features of T2w and DWI images with GGG, but there is a large heterogeneity both in the preselection of quantitative feature to test and in the results. In their study, Nketiah et al. [29] selected 4 distinct textural features extracted from T2w from the 14 GLCM textural features originally proposed by Haralick. They found that among the T2W image textural features ASM and entropy correlated significantly ($p < 0.05$) with GGG. In the study of Wibner et al. [8], higher PZ GGG cancer was associated with higher Entropy and lower Energy extracted from ADC maps. None of the texture features of T2w images showed significant associations with GGG. According to the authors, to effectively predict PCa aggressiveness, the identification of an IF signature could be more effective than the identification of specific and isolated features.

Toivonen et al. [30] have recently proposed a combination of features, selected from a pool of 7105 IFs extracted from T2w, DWI a T2 mapping, in a cohort of patients comparable to ours, showing an AUC of 0.88 for the classification of low vs. high GGG. Considering our cohort dimension and the need of an imaging post-processing methodologies applicable in everyday clinical practice, we aimed to limit the analysis to a reasonable number of accurate and reproducible features. For this reason, in this work, a signature composed of the top 10 stable and non-redundant features was tested.

We also attempted to implement a quantitative model able to predict alongside GGG also ECE and nodal involvement. Knowledge of ECE at the time of diagnosis can affect management decisions, with a potential impact on treatment modality or surgical technique. The preoperative prediction of nodal involvement is even more important for clinical

practice and it is currently based on clinical and pathological data [31]. As a matter of fact, even though some qualitative MRI parameters, such as dimension and length of capsular contact of dominant lesion, have been identified as a predictor of ECE and nodal involvement [32], due to its subjective nature, the evaluation of these parameters is plagued by low inter-observer agreement [33], and a high number of cases still suffer from over diagnosis and overtreatment.

Using a limited combination of IFs from T2w and ADC images, selected among the top 10 more stable features, we were able to predict GGG, ECE, and node status with a sensitivity greater than 0.88 and a specificity of 0.78 at least (Table 3).

Since mean ADC value in tumor volumes is known to be a quite reliable biomarker for tumour aggressiveness [34,35], we tested the effective potential advantage of radiomic IFs with respect to the easier use of ADC_{mean}. Interestingly, mean ADC alone showed a lower accuracy with respect to the use of radiomic IFs, but when used in combination (ADC mean + $S_{TOP\text{-}T2/ADC}$), an improvement in model performance was obtained, particularly for prediction of GGG and ECE (Table 3).

In this study, no information such as genomic and proteomic data of the tumors was included since this information is not usually evaluated in routine clinical practice. However, it would be interesting, in prospective works, to include other biological parameters possibly underlying molecular changes of PC that may reflect in radiomic features (e.g., mRNAs or miRNAs expression levels). Such molecular features could help to decipher the biological role of radiomic features in capturing genotype in its living environment and eventually complete and improve the radiomic signature to predict clinical endpoints.

This study has some limitations most related to the small, even though homogeneous, sample size. Indeed, our inclusion criteria required that enrolled patients performed both mpMRI and image-guided biopsy, thus obtaining a cohort of biopsy-confirmed patients to be analyzed by radiomic mpMRI. Then, patients received prostatectomy as primary treatment, which is nowadays recommended only for patients showing high-grade tumors since can be postponed for localized, low-grade (Gleason < 7) disease without significant change in outcome [36]. Moreover, we also considered a subset of patients undergoing lymphadenectomy during surgery in order to assess the predictiveness of mpMRi-based signatures with respect to lymph node status. Furthermore, in order to warrant rigorous radiomic analysis, we had to exclude, from our enrolled cohort, those patients with lesion dimensions on mMRI below 0.7 cc. We were very conservative on this point to warrant an accurate and meaningful radiomic analysis. This cut-off for lesion volume was due to the lower spatial resolution of ADC maps that were included in our multiparametric signature. In the majority of published radiomics studies on PCa, tumors were analyzed on the basis of their dimensions as evaluated by histology (0.5 cc dimensional cut-off for clinically significant tumors) [8,29,37]. Despite this, a radiomic analysis on less than 4 image voxels in each direction make no meaningful the calculation for different textural parameters [6]. For those reasons, we were not able to perform an external validation of our radiomic signatures without further impacting on samples to be used for training and internal validation.

Limitation in terms of VOI dimensions is clearly an issue of radiomics [6], which could have an impact from a clinical point of view. Despite this, in our cohort with the application of a dimensional cut off of 0.7 cc, the large majority of discarded patients have small PCas on imaging characterized by a low/intermediate aggressiveness at pathology (53% were GGG 3 + 4). As pointed out, the dimensional cut-off was defined considering the lower spatial resolution of ADC maps for our multiparametric signature. However, due to higher resolution of T2w images, the minimum VOI volume was required for radiomic analysis of T2w images. Thus, mono-modal T2w signature could be used to evaluate aggressiveness in case of small PCas.

Even if we did not separate PZ and TZ lesions for the analysis due to the disproportion in frequencies distribution, with PZ tumors accounting for almost 70% of lesions, performances of the different signatures in the subset of PZ lesions have been evaluated. The

obtained results confirmed the predictive role of DWI in PZ lesions, even if the multivariate signature S_{TOP} has shown slightly higher accuracy. Despite this, a larger cohort of TZ lesions will allow to confirm the specific impact of signature in this subset of PC patients. Finally, we focused our investigation on radiomic features extracted from T2w images and ADC map, excluding data from DCE imaging. Although our approach could be considered in agreement with PIRADS guidelines, which consider DCE only as a support to T2 and DWI sequences for lesion detection and characterization, it could be interesting to investigate the additional contribution of radiomic features from DCE.

In conclusion, our data suggest that the implementation of robust radiomic-pipeline for the analysis of mpMRI allows for the extraction of a reliable radiomic signature, potentially able to predict PCa aggressiveness, in terms of Gleason score, extracapsular extension, and nodal stage.

Supplementary Materials: https://www.mdpi.com/2075-4418/11/4/594/s1, Figure S1: ROC curves and AUC values for classification in PZ patient dataset. Table S1: Report on image processing and image biomarker extraction, Table S2: Signature models and performances for classification of PZ lesions.

Author Contributions: Conceptualization, A.D., F.G., and G.C.; methodology, F.G. and C.C.; validation, F.G.; formal analysis, F.G.; investigation, A.D., F.G., G.C., G.B., and M.I.; data curation, A.D., F.G., and G.C.; writing—original draft preparation, A.D., F.G., and G.C.; writing—review and editing, A.E., A.B., F.M., I.C., and F.D.C.; visualization, A.D., F.G., and G.C.; supervision, I.C. and F.D.C.; project administration, I.C. and F.D.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board San Raffaele Hospital (protocol name ProTex and 30/09/2019; ClinicalTrials.gov Identifier NCT04219059).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data sharing not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Walsh, P.C.; Deweese, T.L.; Eisenberger, M.A. Localized Prostate Cancer. *N. Engl. J. Med.* **2007**, *357*, 2696–2705. [CrossRef] [PubMed]
2. Nepple, K.G.; Wahls, T.L.; Hillis, S.L.; Joudi, F.N. Gleason score and laterality concordance between prostate biopsy and prostatectomy specimens. *Int. Braz. J. Urol.* **2009**, *35*, 559–564. [CrossRef] [PubMed]
3. Rajinikanth, A.; Manoharan, M.; Soloway, C.T.; Civantos, F.J.; Soloway, M.S. Trends in Gleason Score: Concordance Between Biopsy and Prostatectomy over 15 Years. *Urology* **2008**, *72*, 177–182. [CrossRef] [PubMed]
4. Steinberg, D.M.; Sauvageot, J.; Piantadosi, S.; Epstein, J.I. Correlation of Prostate Needle Biopsy and Radical Prostatectomy Gleason Grade in Academic and Community Settings. *Am. J. Surg. Pathol.* **1997**, *21*, 566–576. [CrossRef] [PubMed]
5. Castellano, G.; Bonilha, L.; Li, L.; Cendes, F. Texture analysis of medical images. *Clin. Radiol.* **2004**, *59*, 1061–1069. [CrossRef]
6. Castiglioni, I.; Gallivanone, F.; Soda, P.; Avanzo, M.; Stancanello, J.; Aiello, M.; Interlenghi, M.; Salvatore, M. AI-based applications in hybrid imaging: How to build smart and truly multi-parametric decision models for radiomics. *Eur. J. Nucl. Med. Mol. Imaging* **2019**, *46*, 2673–2699. [CrossRef]
7. Chaddad, A.; Niazi, T.; Probst, S.; Bladou, F.; Anidjar, M.; Bahoric, B. Predicting Gleason Score of Prostate Cancer Patients Using Radiomic Analysis. *Front. Oncol.* **2018**, *8*, 630. [CrossRef] [PubMed]
8. Wibmer, A.; Hricak, H.; Gondo, T.; Matsumoto, K.; Veeraraghavan, H.; Fehr, D.; Zheng, J.; Goldman, D.; Moskowitz, C.; Fine, S.W.; 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. *Eur. Radiol.* **2015**, *25*, 2840–2850. [CrossRef]
9. Ms, S.E.V.; Bloch, N.B.; Chappelow, J.C.; Toth, R.; Rofsky, N.M.; Genega, E.M.; Lenkinski, R.E.; Madabhushi, A. Central gland and peripheral zone prostate tumors have significantly different quantitative imaging signatures on 3 tesla endorectal, in vivo T2-weighted MR imagery. *J. Magn. Reson. Imaging* **2012**, *36*, 213–224. [CrossRef]
10. Srigley, J.R.; A Humphrey, P.; Amin, M.B.; Chang, S.S.; Egevad, L.; I Epstein, J.; Grignon, D.J.; McKiernan, J.M.; Montironi, R.; A Renshaw, A.; et al. Protocol for the examination of specimens from patients with carcinoma of the prostate gland. *Arch. Pathol. Lab. Med.* **2009**, *133*, 1568–1576. [CrossRef]

11. Palumbo, D.; Yee, B.; O'Dea, P.; Leedy, S.; Viswanath, S.; Madabhushi, A. Interplay between bias field correction, intensity standardization, and noise filtering for T2-weighted MRI. In Proceedings of the 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Boston, MA, USA, 30 August–3 September 2011; Institute of Electrical and Electronics Engineers (IEEE): Greenvile, SC, USA, 2011; pp. 5080–5083.
12. Tustison, N.J.; Avants, B.B.; Cook, P.A.; Zheng, Y.; Egan, A.; Yushkevich, P.A.; Gee, J.C. N4ITK: Improved N3 Bias Correction. *IEEE Trans. Med. Imaging* **2010**, *29*, 1310–1320. [CrossRef]
13. Fedorov, A.; Beichel, R.; Kalpathy-Cramer, J.; Finet, J.; Fillion-Robin, J.-C.; Pujol, S.; Bauer, C.; Jennings, M.; Fennessy, F.; Sonka, M.; et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn. Reson. Imaging* **2012**, *30*, 1323–1341. [CrossRef] [PubMed]
14. Zhu, L.; Kolesov, L.; Yi, G.; Ron, K.; Allen, T. An Effective Interactive Medical Image Segmentation Method using Fast GrowCut. In Proceedings of the International Conference Medicine Image Comput Comput Assist Interv, Workshop on Interactive Methods (MICCAI), Boston, MA, USA, 14–18 September 2014.
15. Van Griethuysen, J.J.; Fedorov, A.; Parmar, C.; Hosny, A.; Aucoin, N.; Narayan, V.; Beets-Tan, R.G.; Fillion-Robin, J.-C.; Pieper, S.; Aerts, H.J. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res.* **2017**, *77*, e104–e107. [CrossRef] [PubMed]
16. Collewet, G.; Strzelecki, M.; Mariette, F. Influence of MRI acquisition protocols and image intensity normalization methods on texture classification. *Magn. Reson. Imaging* **2004**, *22*, 81–91. [CrossRef]
17. Aerts, H.J.W.L.; Velazquez, E.R.; Leijenaar, R.T.H.; Parmar, C.; Grossmann, P.; Carvalho, S.; Bussink, J.; Monshouwer, R.; Haibe-Kains, B.; Rietveld, D.; et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat. Commun.* **2014**, *5*, 4006. [CrossRef] [PubMed]
18. Gillies, R.J.; Kinahan, P.E.; Hricak, H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* **2016**, *278*, 563–577. [CrossRef]
19. Gu, Z.; Eils, R.; Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **2016**, *32*, 2847–2849. [CrossRef] [PubMed]
20. Bertoli, G.; Cava, C.; Diceglie, C.; Martelli, C.; Rizzo, G.; Piccotti, F.; Ottobrini, L.; Castiglioni, I. MicroRNA-567 dysregulation contributes to carcinogenesis of breast cancer, targeting tumor cell proliferation, and migration. *Breast Cancer Res. Treat.* **2016**, *161*, 605–616. [CrossRef] [PubMed]
21. Meyer, D.; Dimitriadou, E.; Hornik, K.; Weingessel, A.; Leisch, F. Misc functions of the department of statistics, Probability Theory Group In: e1071. *TU Wien. R Package Version1.7–2*, 2019.
22. D'Amico, N.C.; Merone, M.; Sicilia, R.; Cordelli, E.; D'Antoni, F.; Zanetti, I.B.; Valbusa, G.; Grossi, E. Tackling imbalance radiomics in acoustic neuroma. *Int. J. Data Min. Bioinform.* **2019**, *22*, 365. [CrossRef]
23. Haibo, H.; Garcia, E.A. Learning from Imbalanced Data. *IEEE Trans. Knowl. Data Eng.* **2009**, *21*, 1263–1284. [CrossRef]
24. Gallivanone, F.; Panzeri, M.M.; Canevari, C.; Losio, C.; Gianolli, L.; De Cobelli, F.; Castiglioni, I. Biomarkers from in vivo molecular imaging of breast cancer: Pretreatment 18F-FDG PET predicts patient prognosis, and pretreatment DWI-MR predicts response to neoadjuvant chemotherapy. *MAGMA* **2017**, *30*, 359–373. [CrossRef]
25. Ganeshan, B.; Burnand, K.; Young, R.; Chatwin, C.; Miles, K. Dynamic Contrast-Enhanced Texture Analysis of the Liver. *Investig. Radiol.* **2011**, *46*, 160–168. [CrossRef]
26. Teruel, J.R.; Heldahl, M.G.; Goa, P.E.; Pickles, M.; Lundgren, S.; Bathen, T.F.; Gibbs, P. Dynamic contrast-enhanced MRI texture analysis for pretreatment prediction of clinical and pathological response to neoadjuvant chemotherapy in patients with locally advanced breast cancer. *NMR Biomed.* **2014**, *27*, 887–896. [CrossRef] [PubMed]
27. Zwanenburg, A.; Vallières, M.; Abdalah, M.A.; Aerts, H.J.W.L.; Andrearczyk, V.; Apte, A.; Ashrafinia, S.; Bakas, S.; Beukinga, R.J.; Boellaard, R.; et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology* **2020**, *295*, 328–338. [CrossRef] [PubMed]
28. Comelli, A.; Dahiya, N.; Stefano, A.; Vernuccio, F.; Portoghese, M.; Cutaia, G.; Bruno, A.; Salvaggio, G.; Yezzi, A. Deep Learning-Based Methods for Prostate Segmentation in Magnetic Resonance Imaging. *Appl. Sci.* **2021**, *11*, 782. [CrossRef]
29. Nketiah, G.; Elschot, M.; Kim, E.; Teruel, J.R.; Scheenen, T.W.; Bathen, T.F.; Selnæs, K.M. T2-weighted MRI-derived textural features reflect prostate cancer aggressiveness: Preliminary results. *Eur. Radiol.* **2017**, *27*, 3050–3059. [CrossRef]
30. Toivonen, J.; Perez, I.M.; Movahedi, P.; Merisaari, H.; Pesola, M.; Taimen, P.; Boström, P.J.; Pohjankukka, J.; Kiviniemi, A.; Pahikkala, T.; et al. Radiomics and machine learning of multisequence multiparametric prostate MRI: Towards improved non-invasive prostate cancer characterization. *PLoS ONE* **2019**, *14*, e0217702. [CrossRef]
31. Briganti, A.; Larcher, A.; Abdollah, F.; Capitanio, U.; Gallina, A.; Suardi, N.; Bianchi, M.; Sun, M.; Freschi, M.; Salonia, A.; et al. Updated Nomogram Predicting Lymph Node Invasion in Patients with Prostate Cancer Undergoing Extended Pelvic Lymph Node Dissection: The Essential Importance of Percentage of Positive Cores. *Eur. Urol.* **2012**, *61*, 480–487. [CrossRef] [PubMed]
32. Rosenkrantz, A.B.; Shanbhogue, A.K.; Wang, A.; Kong, M.X.; Babb, J.S.; Taneja, S.S. Length of capsular contact for diagnosing extraprostatic extension on prostate MRI: Assessment at an optimal threshold. *J. Magn. Reson. Imaging* **2016**, *43*, 990–997. [CrossRef]
33. Hole, K.H.; Axcrona, K.; Lie, A.K.; Vlatkovic, L.; Geier, O.M.; Brennhovd, B.; Knutstad, K.; Olsen, D.R.; Seierstad, T. Routine pelvic MRI using phased-array coil for detection of extraprostatic tumour extension: Accuracy and clinical significance. *Eur. Radiol.* **2012**, *23*, 1158–1166. [CrossRef]

34. De Cobelli, F.; Ravelli, S.; Esposito, A.; Giganti, F.; Gallina, A.; Montorsi, F.; Del Maschio, A. Apparent Diffusion Coefficient Value and Ratio as Noninvasive Potential Biomarkers to Predict Prostate Cancer Grading: Comparison With Prostate Biopsy and Radical Prostatectomy Specimen. *Am. J. Roentgenol.* **2015**, *204*, 550–557. [CrossRef]
35. Oto, A.; Yang, C.; Kayhan, A.; Tretiakova, M.; Antic, T.; Schmid-Tannwald, C.; Eggener, S.; Karczmar, G.S.; Stadler, W.M. Diffusion-Weighted and Dynamic Contrast-Enhanced MRI of Prostate Cancer: Correlation of Quantitative MR Parameters with Gleason Score and Tumor Angiogenesis. *Am. J. Roentgenol.* **2011**, *197*, 1382–1390. [CrossRef] [PubMed]
36. Gillies, R.J.; Schabath, M.B. Radiomics Improves Cancer Screening and Early Detection. *Cancer Epidemiol. Biomark. Prev.* **2020**, *29*, 2556–2567. [CrossRef] [PubMed]
37. Giambelluca, D.; Cannella, R.; Vernuccio, F.; Comelli, A.; Pavone, A.; Salvaggio, L.; Galia, M.; Midiri, M.; Lagalla, R.; Salvaggio, G. PI-RADS 3 Lesions: Role of Prostate MRI Texture Analysis in the Identification of Prostate Cancer. *Curr. Probl. Diagn. Radiol.* **2021**, *50*, 175–185. [CrossRef] [PubMed]