Elsevier Editorial System(tm) for European Journal of Radiology

Manuscript Draft

Manuscript Number: EJR-D-17-01269R1

Title: Diagnostic accuracy of biparametric vs multiparametric MRI in clinically significant prostate cancer: comparison between readers with different experience.

Article Type: Research Article

Section/Category: Genitourinary Imaging

Keywords: Clinically significant prostate cancer; Prostate Imaging Reporting Data System Version 2; Diffusion weighted imaging; Dynamic contrasted-enhanced imaging; Multi-reader scoring.

Corresponding Author: Dr. Andrea Delli Pizzi,

Corresponding Author's Institution: ITAB Institute of Advanced Biomedical Technologies, University “G. d’Annunzio”

First Author: Eleonora Di Campli

Order of Authors: Eleonora Di Campli; Andrea Delli Pizzi; Barbara Seccia; Roberta Cianci; Martina d'Annibale; Antonella Colasante; Sebastiano Cinalli; Pietro Castellan; Riccardo Navarra; Romina Iantorno; Daniela Gabrielli; Angelica Buffone; Massimo Caulo; Raffaella Basilico

Abstract: Background
MRI plays a crucial role to identify men with a high likelihood of clinically significant prostate cancer who require immediate biopsy. The added value of DCE MRI in combination with T2-weighted imaging and DWI is controversial (risks related to gadolinium administration, duration of MR exam, financial burden, effects on diagnostic performance). A comparison of a biparametric and a standard multiparametric MR imaging protocol, taking into account the different experience of the readers, may help to choose the best MR approach regarding diagnostic performance.
Purpose
To determine the added value of dynamic contrasted-enhanced imaging (DCE) over T2-weighted imaging (T2-WI) and diffusion weighted imaging (DWI) for the detection of clinically significant prostate cancer, and to evaluate how it affects the diagnostic performance of three readers with different grade of experience in prostate imaging.
Materials and Methods
Eighty-five patients underwent prostate MR examination at 1.5T MR scanner performed because of elevated prostate-specific antigen level and/or suspicion of prostate cancer at digital rectal examination.
Two MR images sets (Set 1=biparametric, Set 2=multiparametric) were retrospectively and independently scored by three radiologists with 7, 3 and 1 years of experience in prostate MR imaging respectively, according to PI-RADS v2. Sensitivity, specificity, positive predictive value, and negative predictive value were calculated by dichotomizing reader scores. Receiver operating characteristic (ROC) analysis was performed and areas under the curve (AUCs) were calculated for each reader and image set. A

comparison of ROC curves was performed to test the difference between the areas under the ROC curves among the three readers.
Results
There was no significant difference regarding the detection of clinically significant tumor among the three readers between the two image sets. The AUC for the bi-parametric and multi-parametric MR imaging protocol was respectively 0.68-0.72 (Reader 1), 0.72-0.70 (Reader 2) and 0.60-0.54 (Reader 3). ROC curve comparison revealed no statistically significant differences for each protocol among the most experienced (Reader 1) and the other readers (Readers 2-3).
Conclusion
The diagnostic accuracy of a bi-parametric MR imaging protocol consisting of T2-weighted imaging and DWI is comparable with that of a standard multi-parametric imaging protocol for the detection of clinically significant prostate cancer. The experience of the reader does not significantly modify the diagnostic performance of both MR protocols.

***Title Page (including paper title & Complete corresponding author details)**

TITLE:

Diagnostic accuracy of biparametric vs multiparametric MRI in clinically significant prostate cancer: comparison between readers with different experience.

AUTHORS:

Eleonora Di Campli[1;]

Andrea Delli Pizzi[1,2];

Barbara Seccia[1];

Roberta Cianci[1];

Martina d'Annibale[1];

Antonella Colasante[3];

Sebastiano Cinalli[4];

Pietro Castellan[5];

Riccardo Navarra[2];

Romina Iantorno[4];

Daniela Gabrielli[1];

Angelica Buffone;

Massimo Caulo[1,2];

Raffaella Basilico[1].

AFFILIATIONS:

1. Department of Neuroscience, Imaging and Clinical Sciences, University "G. d'Annunzio", Chieti, Italy
2. ITAB Institute of Advanced Biomedical Technologies, University "G. d'Annunzio", Via Luigi Polacchi, 11 66100 Chieti, Italy
3. Surgical Pathology Unit, "SS Annunziata" Hospital, ASL2 Abruzzo, Chieti, Italy.
4. Department of Medicine and Aging Sciences, Section of Pathological Anatomy, G. D'Annunzio University.
5. Department of Urology, ASL02 Abruzzo, Via dei Vestini, 66100, Chieti, Italy.

CORRESPONDING AUTHOR

Andrea Delli Pizzi

andreadellipizzi@gmail.com

Institute of Advanced Biomedical Technologies, University "G. d'Annunzio", Via Luigi Polacchi, 11 66100 Chieti, Italy

08713556901

ORCID: 0000-0002-2011-3753

ABSTRACT:

Background

MRI plays a crucial role to identify men with a high likelihood of clinically significant prostate cancer who require immediate biopsy. The added value of DCE MRI in combination with T2-weighted imaging and DWI is controversial (risks related to gadolinium administration, duration of MR exam, financial burden, effects on diagnostic performance). A comparison of a biparametric and a standard multiparametric MR imaging protocol, taking into account the different experience of the readers, may help to choose the best MR approach regarding diagnostic performance.

Purpose

To determine the added value of dynamic contrasted-enhanced imaging (DCE) over T2-weighted imaging (T2-WI) and diffusion weighted imaging (DWI) for the detection of clinically significant prostate cancer, and to evaluate how it affects the diagnostic performance of three readers with different grade of experience in prostate imaging.

Materials and Methods

Eighty-five patients underwent prostate MR examination at 1.5T MR scanner performed because of elevated prostate-specific antigen level and/or suspicion of prostate cancer at digital rectal examination.

Two MR images sets (Set 1=biparametric, Set 2=multiparametric) were retrospectively and independently scored by three radiologists with 7, 3 and 1 years of experience in prostate MR imaging respectively, according to PI-RADS v2. Sensitivity, specificity, positive predictive value, and negative predictive value were calculated by dichotomizing reader scores.

Receiver operating characteristic (ROC) analysis was performed and areas under the curve (AUCs) were calculated for each reader and image set. A comparison of ROC curves was performed to test the difference between the areas under the ROC curves among the three readers.

Results

There was no significant difference regarding the detection of clinically significant tumor among the three readers between the two image sets. The AUC for the bi-parametric and multi-parametric MR imaging protocol was respectively 0.68-0.72 (Reader 1), 0.72-0.70 (Reader 2) and 0.60-0.54 (Reader 3). ROC curve comparison revealed no statistically significant differences for each protocol among the most experienced (Reader 1) and the other readers (Readers 2-3).

Conclusion

The diagnostic accuracy of a bi-parametric MR imaging protocol consisting of T2-weighted imaging and DWI is comparable with that of a standard multi-parametric imaging protocol for the detection of clinically significant prostate cancer. The experience of the reader does not significantly modify the diagnostic performance of both MR protocols.

KEYWORDS

Clinically significant prostate cancer; Prostate Imaging Reporting Data System Version 2; Diffusion weighted imaging; Dynamic contrasted-enhanced imaging; Multi-reader scoring.

Author Contributions All authors were involved in patient management and wrote the report. Written consent to publication was obtained.

Compliance with Ethical Standards

Conflict of Interest All authors declare that they have no conflict of interest.

Ethical Approval All procedures performed in this study involving human participant were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent from patients included in this retrospective study was waived.

INTRODUCTION:

Prostate cancer is the most frequently diagnosed cancer in males and the second leading cause of cancer-related death in men [1]. Multiparametric magnetic resonance imaging (MRI) has become an important tool for the diagnosis of prostate cancer. It is a prerequisite for optimal clinical management and therapy selection because it is particularly helpful in the detection of prostate cancer foci, local staging, and the estimation of prostate cancer aggressiveness [2-4].

Prostate MRI is increasingly performed before a biopsy in patients with elevated PSA. It is used as an additional parameter all together with digital rectal examination and patient's age and comorbidity to identify men with a high likelihood of clinically significant prostate cancer requiring immediate biopsy, because of its potential for causing death. Prostate MRI may also improve the biopsy yield by targeting suspicious lesions, while reducing the risk of unnecessary diagnosis of clinically insignificant tumors.

The Prostate Imaging Reporting Data System (PI-RADS), published in 2012 by the European Society of Urogenital Radiology (ESUR), recommended to include multiplanar T1- and T2-weighted images, diffusion-weighted imaging (DWI) and dynamic contrast-enhanced (DCE) MRI, in a multiparametric MRI protocol [5].

Because of non-conclusive results concerning the value of DCE MRI in the detection of prostate cancer, particularly for evaluation of the transitional zone, the added value of DCE MRI in combination with T2-weighted imaging and DWI is still under debate [6-9].

The role of DCE MRI was recently downgraded by the American College of Radiology and the European Society of Urogenital Radiology that provided the new Prostate Imaging Reporting and Data System (PI-RADS) version 2. In detail, DWI for the peripheral zone (PZ) and T2-WI for the transient zone (TZ) were considered the two dominant sequences to differentiate non-clinically significant and clinically significant tumors. In this way, the role of DCE is minor considering the peripheral zone (PZ) and even unuseful concerning the TZ for the differentiation between prostate cancer and benign prostatic hyperplasia [10]. In this context, MR scan time, patient comfort and costs could beneficiate of a protocol devoid of DCE; moreover, potential risks related to the use of intravenous contrast such as nephrogenic systemic fibrosis, renal failure and brain accumulation of gadolinium would be reduced [11].

Thus, the purpose of this study was to determine the added value of dynamic contrasted-enhanced imaging (DCE) over T2-weighted imaging (T2-WI) and diffusion weighted imaging (DWI) using a 1.5 T scanner without endorectal coil for the detection of clinically significant prostate cancer. Moreover, we wanted to evaluate how these two different protocols affect the diagnostic performance of three readers with different grade of experience in prostate imaging.

MATERIALS & METHODS:

All procedures performed in this study involving human participant were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. A total of 160 patients underwent multi-parametric-MRI (mp-MRI) of the prostate between March 2013 and December 2016 were retrospectively included; informed consent was waived. These patients were identified through our RIS/PACS (Radiology Information System/Picture Archiving and Communication System).

Inclusion criteria

- Multiparametric MRI of prostate at 1.5T, without endorectal coil, performed because of elevated prostate-specific antigen level and/or suspicion of prostate cancer at digital rectal examination;
- standardized prostatic biopsy performed within 6 months pre/after MR imaging examination e/o prostatectomy performed within 6 months after MR imaging examination;
- negative follow-up greater than 6 months (considering PSA level and digital rectum examination) in case of negative biopsy;
- adenocarcinoma type tumors in case of positive biopsy.

Exclusion criteria

- not available (patient underwent surgery in other hospitals)/inconclusive histological examination;
- not satisfactory images (multiple artifacts from, for example, total hip replacements, patient movements);
- hormonal therapy before MR imaging examination.

The final study population comprised 85 patients (mean age, 70,39 years; range, 54-84 years) with a mean prostate-specific antigen level of 8,5 mg/L (range, 0,77- 27,74 mg/L; follow up range 8-36 months) (Figure 1). A total of 42 biopsies and 43 prostatectomies were included. No patietis were dropped out due to the lack of follow-up information.

MRI protocol

MR exams were performed using a 1.5-T scanner (Achieva; Philips Medical System; Best; Netherlands), without endorectal coil. MR protocol included T1-weighted (T1W) fast spin-echo, without fat suppression, images, T2-weighted (T2W) turbo spin-echo images, diffusion-weighted images (DWI) and dynamic contrast-enhanced (DCE) T1-weighted (T1W) 3D spoiled gradient-echo images (Table 1). Apparent diffusion coefficient maps (ADC-maps) were calculated for each patient. Gadobenate dimeglumine (Multihance; Bracco, Milan, Italy) was used in a dose of 0.1 mmol/kg of body weight (flow rate of 2 mL/sec). If there were no contraindications, intravenous injection of hyoscinbutylbromid (Buscopan, 20mg/mL, injection fluid, Boehringer, Ingelheim, Germany) was administered to reduce peristaltic movement.

Image Analysis

The final study population included 85 patients for a total of 85 MR scans (each one including T2w, DWI and DCE). For the reading-session, MR scans were organised in two sets of images, Image set 1 and Image set 2.

Image set 1 comprised the bi-parametric MRI protocol. It was presented in the first reading session, and consisted of:

Comment [ADP1]: Reviewer 1
2)The study is of retrospective nature. There is a numerical imbalance between patients with prostate cancer and patients without detected cancer poining to an existing bias. Thus, patient selection may be confounded by factors such as patient compliance with follow-up examinations or biopsy appointments.

The reviewer is right in that the retrospective nature is a weakness of the study. We considered it as a limit and in the manuscript (page 7 – line 21-22) we specified that "our results need to be further validated, possibly with multicentric and prospective design studies". The numerical imbalance between patients with prostate cancer and patients without cancer reflects the fact that in our study population, most patients underwent MR after a positive biopsy. All biopsies were performed within 6 months pre/after MR and there were no bias of selection due to the lack of follow-up information.
A statement was included in methods: "No patients were dropped out due to the lack of follow-up information."

Comment [ADP2]: Reviewer 2
2) It is not clearly stated if there were 85 MRIs in each Set or if two groups were formed with each group containing half of the study population. When discussing the results it would also help to state some percentages to put numbers into perspective.

A statement was included to clarify this point according tot he reviewer's suggestion.

a) axial T1-weighted and axial, sagittal and coronal T2-weighted images;

b) axial DWI images (including the corresponding ADC maps).

Image set 2 was evaluated in the second reading session and included the complete protocol (T1-weighted images, T2-weighted images, DWI and DCE-MRI).

Two image sets (Image set 1 and Image set 2) were independently reviewed by three radiologists with different grade of experience in interpreting prostate MR imaging (7 years Reader 1; 3 years Reader 2; 1 years Reader 3), in two separate reading sessions.

MRI criteria for malignancy were assessed according to PI-RADS v2 as well as the prostate subdivision in sectors [10]. In this way, thirty-nine sectors (thirty-six for the prostate, two for the seminal vesicles and one for the external urethral sphincter) were considered. Following PI-RADS v2, all detected lesions were scored from 1 to 5 and, more in detail, 1 corresponded to a very low probability for clinically significant prostatic carcinoma; 2 identified low probability; 3 was used for intermediate probability; 4 was associated with high probability; 5 was related to very high probability). According to Weinreb et al., when no contrast enhanced data were available (image set 1), DWI sequences were of key importance for the final PI-RADS score if the tumor was in the Peripheral Zone. Otherwise, both T2-weighted and DWI sequences were used to calculate the final PI-RADS score in case of tumors in the Transitional Zone [12].

Comment [ADP3]: Reviewer 2
2) It is not clearly stated if there were 85 MRIs in each Set or if two groups were formed with each group containing half of the study population. When discussing the results it would also help to state some percentages to put numbers into perspective.

This issue was moved after the description of the two image sets.

Comment [ADP4]: Reviewer 1
6) How was PIRADS used when no contrast-enhanced data were available?
In absence of contrast-enhanced scans, a PIRADS Assessment Category "X" was assigned, according to PIRADS v2. In detail, concerning Peripheral Zone, PIRADS score was obteind by DWI Assessment Cathegory, since DWI are of key importance in that zone. Regarding Transition Zone, both T2 weighted and DWI sequences were used to calculate the PIRADS score. In the main text a statement was included ("Image Analysis"): "According to Weinreb ed al., when no contrast-enhanced data were available (image set 1) considering the bi-parametric MRI protocol, DWI sequences were of key importance for the final PI-RADS score if the tumor was in the Peripheral Zone. Otherwise, both T2-weighted and DWI sequences were used to calculate the final PI-RADS score in case of tumors in the Transitional Zone". A reference was added in the manuscript.

Weinreb JC et al., PI-RADS Prostate Imaging – Reporting and Data System: 2015, Version 2. Eur. Urol.2016 Jan;69(1):16-40.

Standard of Reference

All patients underwent biopsy and a dedicated genitourinary pathologist assessed histopathologic specimens before or after multi-parametric MR imaging or prostatectomy within 6 months (mean, +/- 78.26 days). A Gleason score ≥7 (3+4) and/or as the presence of extra-prostatic invasion defined the presence of clinically significant cancer. A fifth-year radiology resident and a third-year pathology resident, both not involved in the readings, matched the information from MR images and histopathologic specimens of biopsies and/or prostatectomy in order to confirm the spatial correspondence of the highest PI-RADS score region with the highest Gleason Score, considered as standard reference.

Statistical Analysis

Diagnostic performance regarding sensitivity, specificity, positive predictive value, and negative predictive value were calculated by dichotomizing reader scores and by using descriptive statistics. In detail, regarding the overall accuracy for the detection of tumor, regardless the clinical significance, PI-RADS v2 scores of 1–2 were considered negative, while PI-RADS v2 scores of 3–4–5 were considered positive. Concerning the accuracy for the detection of clinically significant prostate cancer, PI-RADS v2 scores of 1–2–3 were considered negative, while PI-RADS v2 scores of 4–5 were considered positive.

Receiver operating characteristic (ROC) analysis was performed and areas under the curve (AUCs) were calculated for each reader and image set. A comparison of ROC curves was performed to test the difference between the areas under the ROC

curves among the three readers. ROC curve comparison was performed with MedCalc software, version 16.8.4 (MedCalc Software, Ostend, Belgium). All other statistical analyses were performed by using IBM SPSS Statistics software, version 20 (IBM, Armonk, NY). A p-value ≤0.05 was considered statistically significant.

RESULTS:

Out of 85 patients, 72 were positive for prostate cancer to biopsy and/or prostatectomy and 13 were negative for cancer. Among 72 tumors, 41 were clinically significant (29 patients with a Gleason 7, 6 patients with a Gleason 8 and 6 patients with Gleason 9) while 31 were not clinically significant (24 Gleason 6, 3 Gleason 5 and 4 Gleason 4). No patients developed a clinically significant cancer during the follow-up.

The comparison of ROC curves among the three readers regarding the overall cancer detection and clinically significant cancer detection using the two different image protocols is shown in Figure 2.

When considering clinically significant prostate cancer, there were no statistically significant differences in diagnostic performance for all the three readers between the two protocols. In fact, the AUC for the bi-parametric and multi-parametric MR imaging protocol was respectively 0.68-0.72 (Reader 1), 0.72-0.70 (Reader 2) and 0.60-0.54 (Reader 3). If we consider reader's experience there were no statistically significant differences in diagnostic accuracy for each protocol among the most experienced (Reader 1) and the others (Readers 2-3).

When considering the overall detection of prostate cancer, there were no differences in diagnostic accuracy for each reader between the two protocols. The AUC for the bi-parametric and multi-parametric MR imaging protocol was respectively 0.79-0.79 (Reader 1), 0.56-0.56 (Reader 2) and 0.64-0.64 (Reader 3). In that case, if we consider readers' experience, AUC of the most experienced reader (Reader 1) was higher compared to the others but the difference was significant only between Reader 1 and Reader 2, for both imaging protocols (p=0.001 Set 1, p=0.001 Set 2) (Table 2).

Diagnostic performance of the readers according to descriptive statistics, ROC analysis and Gleason score are shown in Table 2, 3 and 4 (Supplementary Material) respectively. An example of correctly classified clinically significant tumor is shown in Figure 3.

DISCUSSION:

Our results demonstrate that a biparametric protocol for MR imaging of the prostate has the same diagnostic accuracy of multiparametric MRI, in the evaluation of clinically significant prostate cancer, regardless the reader's experience.

The increasing interest for the detection of clinically significant cancer lead researchers to test different MR protocols but results are currently inconclusive. In fact, a recent systematic review by Fütterer JJ et al. revealed not negligible variations in terms of MRI accuracy (range 4%-87%), sensitivity (58%-96%) and specificity (23%-87%) [13].

The reasons for these variations may be twofold. Firstly, the five-point scale used to score the probability of the clinically significant tumor being present needs to be dichotomized for ROC analysis. In this regard, taking into account that a score

Comment [ADP5]: Reviewer 1
5) Methods: Available histopathologic correlation is heterogenious in this study including biopsies as well as prostatectomy specimens. It is thus difficult to determine the NPV and specificity of the MR exams.

The reviewer is right when saying that the histopathologic correlation is heterogenious and this is a potential weakness of the study. In fact, despite of the high number of prostatectomies (43) and the long term follow up (> 6 months), the availability of the biopsy alone as gold standard for certain patients may represent a limitation because they were performed "random" and not targeted. In this way, if on one hand a negative follow up after a negative biopsy (all patients included in our study showed negative follow-up) is certainly comfortable, on the other hand it is not always possible to exclude the presence of subclinical cancers. For this reason, in case of false negative biopsy, this could have underestimated the presence of prostate cancer and, consequently, it could be responsible for underestimation of the diagnostic performance of the readers as well. This point was discussed in the discussion/limitations.
In the results was added the statement "No patients developed a clinically significant cancer during follow-up".

Comment [ADP6]: Reviewer 1
8) The number of figures and tables can be reduced presenting the most important aspects. A supplement should be perhaps considered.

The Table 4 was moved in "supplements" as suggested by the reviewer.

Comment [ADP7]: Corrected

of 3 refers to a low probability of clinically significant cancer, there is no relevant agreement among the studies concerning if it would be considered a positive or a negative MR finding. In our study, a score of 3 was considered as negative regarding the presence of clinically significant cancer, whereas it was considered as positive for the assessment of the overall cancer diagnostic accuracy. Secondly, although the definition of clinically significant cancer includes tumors with a volume greater than 0.5 ml regardless the Gleason score, several studies do not consider any volumetric criteria to define the clinically significant cancer patients group, as in our study. Therefore, in this kind of studies the number of tumor foci considered for the analysis could have been higher than those using this volumetric criterion, leading to conflicting results.

Taking into account these factors, the best overall accuracy achieved in our study was 0.79 while the best clinically significant cancer accuracy was 0.72 with no significant differences between the two MR protocols. Although these performances seem to be in line with results of earlier reports [19], they are not optimal and do not beneficiate of the use of DCE-MRI. A possible reason can be found in the fact that in our study we had a high number of tumors with a Gleason score = 6 (24/85 patients, 28.2%). These tumors are considered not clinically significant but were responsible for the most of the false positive cases. In detail, the expert reader scored them as 4 or 5 according to PIRADS v2 in nine cases using a biparametric protocol and in ten cases using a multiparametric protocol. Furthermore, a range of 10-19 false positive cases was recorded for the two less expert readers. In fact, these tumors can show no differences in MR signal characteristics compared with Gleason 7 cancers because of their similar conspicuity on DWI and DCE-MRI [14-16]. Moreover, several other pitfalls such as physiological changes in the peripheral and central zone, stromal benign prostatic hyperplasia, acute/chronic prostatitis, may have hampered the diagnostic performance of less experienced readers [17, 18].

The limited added value of DCE MR imaging in a multi-parametric protocol is confirmed by our data, which showed similar AUCs with or without the use of DCE MR imaging. In addition, DCE images might have confused the less experienced reader because contrast enhancement in benign prostate hyperplasia can mimic prostate cancer and may, therefore, have led to false-positive results, while at the same time low-grade tumors can show flow kinetics similar to those of normal tissue, potentially leading to underestimation of tumor foci. An interesting finding of our study is that while the overall detection of prostate cancer is influenced by reader's experience, no significant differences were shown regarding the detection of clinically significant prostate between the most experienced reader and the Reader 2 and Reader 3. This could be due to PI-RADS classification, which standardizes the reporting, aiding less experienced readers.

The question whether the biparametric and multiparametric protocols were comparable in terms of clinically significant prostate cancer is not novel and many studies support our results. For example, Radtke and al. investigated the usefulness of biparametric MR for the detection of prostate cancer in the anterior fibromuscular stroma and the transition zone of the prostate. They demonstrated that biparametric MR is not inferior to multiparametric MR in terms of detection of clinically significant cancer with the advantage to be more cost efficient [20]. The study of Stanzione et al., published in 2016, revealed that multiparametric MR is not superior to biparametric MR in the detection of clinically significant prostate cancer, as well [21]. Khul et al. supported these results proposing a short-time MR protocol for patients with elevated prostate specific antigen [22]. In this context, in the light of the increased number of MR exams, researchers investigated new methods to reduce as much as possible the scan-time. Weiss et al. recently assessed the feasibility of simultaneous multislice echo-planar DWI and compared their quality with that of single-shot echo-planar DWI. They showed a

Comment [ADP8]: Reviewer 2
3)Regarding the literature cited in this article I would include some more studies on the topic of biparametric MRI for the detection of significant prostate cancer in the discussions of the results to put their own results into context and to support the positive outcome of there study.

Authors: accordingly to the reviewer's suggestion, three recent studies were included.

substantial reduction of the scan time maintaining a similar image quality [23].

However, according to our knowledge, no previous studies considering readers of different grade of experience in assessing multi-parametric MR images of the prostate are present in literature. Most of previous studies consider all expert readers with several years of experience [9, 19, 21]. Therefore, our results might be more generalizable and applicable to less-experienced radiologists. Furthermore, considering that MRI of prostate is now a very important useful tool not only for cancer detection but also for cancer treatment once the diagnosis has been made, such as active surveillance, the results of our study may offer benefits in terms of cost efficiency, scan time and patient safe.

Our study had some limitations. First, a potential limitation concerns the use of a field strength of 1.5 T without endorectal coil instead of 3T. In spite of this, our study results are in line with those of previous studies performed using a 3T MR scanner [9, 21, 24]. Moreover, credible results were obtained with 1.5T without use of endorectal coil on condition that radiologist supervise and optimize protocols to obtain the best image quality possible [12, 25]. Second, although the high number of prostatectomy (43) and the long term follow up (> 6 months), the availability of the biopsy alone as gold standard for certain patients may represent a limitation because they were performed "random" and not targeted. In this way, if on one hand a negative follow up after a negative biopsy (all patients included in our study showed negative follow-up) is certainly comfortable, on the other hand it is not always possible to exclude the presence of subclinical cancers. For this reason, in case of false negative biopsy, this could have underestimated the presence of prostate cancer and, consequently, it could be responsible for underestimation of the diagnostic performance of the readers as well. Third, we included Gleason 4 and 5 under "not clinically significant" tumors. Although this choice does not impact neither with the results concerning the clinically significant cancer detection (they are considered in the same group of the Gleason 6), neither with the results of the overall cancer detection (there were no significant differences between the biparametric and the multiparametric approach), it should be considered as a limitation. In fact, in recent years Gleason scores of 4 and 5 have mostly disappeared as pointed out by the 2014 International Society of Urological Pathology Consensus Conference [26]. Fourth, due to the retrospective nature of our study, our results need to be further validated, possibly with multicentric and prospective design studies.

CONCLUSION:

The diagnostic accuracy of a bi-parametric MR imaging protocol consisting of T2-weighted imaging and DWI is comparable with a standard multi-parametric imaging protocol for the detection of clinically significant prostate cancer. Moreover, the experience of the reader does not affect the detection of clinically significant prostate cancer by using biparametric or multiparametric MRI protocols and PI-RADS V2 classification. Further studies are necessary to assess if the DCE MR imaging may definitively be omitted from the standard protocol or only be acquired for specific indications.

Comment [ADP9]: Reviewer 1
4) Recent publications proposing short protocols for prostate cancer MR imaging should be included in the introduction (e.g. Barth et al 2017, Weiß et al . 2017).

Authors: Accordingly to the reviewer's suggestion, we added the two references and a short statement.

Comment [ADP10]: Reviewer 1
1)The scope and results of this study are not novel. It is already well-estalished that the main diagnostic information if prostate MRI lies in the T2w and t DWI sequences. This already reflected in the currect version of PIRADS and will likely be even more pronounced in future versions. ...

Comment [ADP11]: Reviewer 1
7) Please discuss the limitations of this paper in more detail in the discussion.

A detailed discussion was now developed according to the reviewer suggestion.

Comment [ADP12]: Reviewer 1
3) The protocol parameters used in this study are not within the recommendations of the PIRADS system. Thus, it is questionable how the presented results can be generalized to alternative study ...

Comment [ADP13]: Reviewer 1
5) Methods: Available histopathologic correlation is heterogenious in this study including biopsies as well as prostatectomy specimens. It is thus difficult to determine the NPV and specificity of the MR exams. ...

Comment [ADP14]: Reviewer 2
3) The authors also include a Gleason Score of 4 and 5 under "not clinically significant" tumors. I would suggest talking to a pathologist about the Gleason scoring system. In recent years Gleason scores of 4 and 5 have mostly disappeared. So the question ...

Comment [ADP15]: Reviewer 2
1)A weak point of this study is probably that the data was analysed retroperspectively.

Authors: The reviewer is right. We considered it as a limitation and in the manuscript, we specified that " ...

REFERENCES:

Comment [ADP16]: References were updated

1. American Cancer Society. Prostate cancer facts. In: Cancer facts & figures 2009. Atlanta, Ga: American Cancer Society, 2009; 19–20.
2. Scheenen TW, et al., Multiparametric magnetic resonance imaging in prostate cancer management: current status and future perspectives. Invest Radiol 2015;50(9):594–600.
3. Rosenkrantz AB, et al., Length of capsular contact for diagnosis of extra-prostatic extension on prostate MRI: Assessment at an optimal threshold. J Magn Reson Imaging 2016;43(4):990–997.
4. Rosenkrantz AB, et al., 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 2015;41(3):708–714.
5. Barentsz JO, Richenberg J, Clements R, et al. ESUR prostate MR guidelines 2012. Eur Radiol 2012;22(4):746–757.
6. Tanimoto A, et al., Prostate cancer screening: the clinical value of diffusion weighted imaging and dynamic MR imaging in combination with T2-weighted imaging. J Magn Reson Imaging 2007;25(1): 146–152.
7. Tamada T, et al., Prostate cancer detection in patients with total serum prostate-specific antigen levels of 4-10 ng/mL: diagnostic efficacy of diffusion weighted imaging, dynamic contrast-enhanced MRI, and T2-weighted imaging. AJR Am J Roentgenol 2011;197(3):664–670.
8. Schimmöller L, et al., MR-sequences for prostate cancer diagnostics: validation based on the PI-RADS scoring system and targeted MR-guided in-bore biopsy. Eur Radiol 2014;24(10):2582–2589.
9. Thestrup KC, et al., Biparametric versus multiparametric MRI in the diagnosis of prostate cancer. Acta Radiol Open 2016 Aug 17;5(8):2058460116663046. doi: 10.1177/2058460116663046. eCollection 2016 Aug.
10. Vargas HA, et al., Updated prostate imaging reporting and data system (PIRADS v2) recommendations for the detection of clinically significant prostate cancer using multiparametric MRI: critical evaluation using whole-mount pathology as standard of reference. Eur Radiol 2016;26(6):1606–1612.
11. Borna K. Barth, et al., Detection of clinically significant Prostate cancer: Short Dual– Pulse Sequence versus Standard Multiparametric MR Imaging—A Multireader Study. Radiology 2017; 000:1–12.
12. Weinreb JC et al., PI-RADS Prostate Imaging – Reporting and Data System: 2015, Version 2. Eur. Urol.2016 Jan;69(1):16-40.
13. Fütterer JJ, et al., Can clinically significant prostate cancer be detected with multiparametric magnetic resonance imaging? A systematic review of the literature. Eur Urol 2015;68(6):1045–1053.
14. Doo KW, et al., Detectability of low and intermediate or high risk prostate cancer with combined T2-weighted and diffusion-weighted MRI. Eur Radiol. 2012 Aug;22(8):1812-9. doi: 10.1007/s00330-012-2430-5. Epub 2012 Apr 1.
15. Verma S, et al., Assessment of aggressiveness of prostate cancer: correlation of apparent diffusion coefficient with histologic grade after radical prostatectomy. AJR Am J Roentgenol. 2011 Feb;196(2):374-81. doi: 10.2214/AJR.10.4441.

16. Tamada T, et al., High b value (2,000 s/mm2) diffusion weighted magnetic resonance imaging in prostate cancer at 3Tesla: comparison with 1,000 s/mm2 for tumor conspicuity and discrimination of aggressive ness. PLoS One. 2014 May 6;9(5):e96619. doi: 10.1371/journal.pone.0096619. eCollection 2014.
17. Quon JS, et al., False positive and false negative diagnoses of prostate cancer at multiparametric prostate MRI in active surveillance. Insights Imaging 2015 Aug;6(4):449-63. doi: 10.1007/s13244-015-0411-3. Epub 2015 May 23.
18. Panebianco V, et al., Pitfalls in Interpreting mp-MRI of the Prostate: A Pictorial Review with Pathologic Correlation. Insights Imaging 2015 Dec;6(6):611-30. doi: 10.1007/s13244-015-0426-9. Epub 2015 Sep 18.
19. Barth BK, et al., Detection of Clinically Significant Prostate Cancer: Short Dual– Pulse Sequence versus Standard Multiparametric MR Imaging—A Multireader Study. Radiology 2017 Mar 27:162020. doi: 10.1148/radiol.2017162020.
20. Radtke JP, et al., Improved detection of anterior fibromuscular stroma and transition zone prostate cancer using biparametric and multiparametric MRI with MRI-targeted biopsy and MRI-US fusion guidance. Prostate Cancer Prostatic Dis. 2015 Sep;18(3):288-96.
21. Stanzione A, et al., Biparametric 3t magnetic resonance imaging for prostatic cancer detection in a biopsy-naïve patient population: A further improvement of PI-RADS v2?. Eur. J. Radiol. 85 (12) (2016) 2269–2274.
22. Kuhl CK, et al., Abbreviated Biparametric Prostate MR Imaging in Men with Elevated Prostate-specific Antigen. Radiology 2017 Nov;285(2):493-505.
23. Weiss J, et al., Feasibility of accelerated simultaneous multislice diffusion-weighted MRI of the prostate. J Magn Reson Imaging 2017 Nov;46(5):1507-1515.
24. De Visschere P, et al., Dynamic contrast-enhanced imaging has limited added value over T2-weighted imaging and diffusion-weighted imaging when using PI-RADSv2 for diagnosis of clinically significant prostate cancer in patients with elevated PSA, Clinical Radiology (2016), http://dx.doi.org/10.1016/j.crad.2016.09.011.
25. Lee SH et al., Is endorectal coil necessary for the staging of clinically localized prostate cancer? Comparison of non-endorectal versus endorectal MR imaging. World J Urol. 2010 Dec;28(6):667-72.
26. Epstein JI, 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. Am J Surg Pathol. 2016 Feb;40(2):244-52.

Comment [ADP17]: Reviewer 1 Recent publications proposing short protocols for prostate cancer MR imaging should be included in the introduction (e.g. Barth et al 2017, Weiß et al . 2017).

Figure 1. Flowchart of patients considered for inclusion, excluded patients and patients finally included in the study cohort.

Figure 2. Comparison of ROC curves among the three readers regarding the overall cancer detection and clinically significant cancer detection using the two image sets (biparametric MRI and multiparametric MRI).

Figure 3. Example of correctly classified clinically significant tumor (Gleason Score 7 proven by radical prostatectomy). A focal hypointensity is visible in the right posterior peripheral zone on T2-weighted MRI. The lesion is markedly hyperintense on DWI and hypointense on ADC map and all the readers assigned a PI-RADS score of 4 for the Set 1 of MR images. DCE-MRI revealed an early and clear enhancement and did not changed the opinion of the readers in the Set 2 of MR images. Radical prostatectomy (RP) confirmed the presence of acinar adenocarcinoma (red arrow) with Gleason Score 7 (4+3) in the right posterior peripheral zone.

Figure 4. Example of uncorrectly classified non-clinically significant tumor (biopsy proven Gleason Score 6). A focal hypointensity is visible in the left posterior peripheral zone on T2-weighted MRI. The lesion is mild hyperintense on DWI and moderately hypointense on ADC map and the expert reader (Reader 1) assigned a PI-RADS score of 3 for the Set 1 of MR images. DCE-MRI revealed an early and clear enhancement by inducing the Reader 1 to increase the Gleason score to 4 for the Set 2 of MR images. The less expert readers (Reader 2 and 3) assigned a PI-RADS score of 4 for the Set 1 of images by interpreting the hyperintense signal on DWI and the hypointense signal on ADC map as focal and marked. They confirmed the PIRADS 4 score at DCE-MRI.

Table 1. Sequence parameters of the T1-weighted, T2-weighted, DWI and DCE sequences used during the study.

	DWI*	**T2-weighted FSE**	**T1-weighted FSE**	**DCE‡**
Image Set 1	Yes	Yes	Yes	No
Image Set 2	Yes	Yes	Yes	Yes
Repetition Time (msec)	2774	7000	700	7.6
Echo Time (msec)	55	120	15	3.7
In-plane resolution (mm)	2x2	0.8x1	0.8x1	1x1
Section thickness (mm)	3	3	5	3
Imaging planes	Transverse†	Transverse†,Coronal, Sagittal	Transverse†	Transverse†
Acquisition time (min)	3.3	4.5	2.5	6.0
Total acquisition time Set 1 (min)	10.3	10.3	10.3	10.3
Total acquisition time Set 2 (min)	16.3	16.3	16.3	16.3

Note: T2-weighted FSE sequences included ADC map calculation. T1-weighted FSE sequences included fat saturation.
* DWI performed with b values of 0, 150, 500, and 1000 sec/mm.
† Transverse plane angulated perpendicular to long axis of gland; DWI and T2-W imaging performed with field of view encompassing prostate gland and seminal vesicles; T1-W imaging performed with field of view encompassing whole pelvis.
‡ DCE performed with temporal resolution of < 8 seconds.

Table 2. Diagnostic performance of the 3 readers regarding the overall cancer detection and the clinically significant cancer detection for image Set 1 and image Set 2.

		Overall cancer detection				**Clinically significant cancer detection**			
		Sensitivity (CI 95%)	**Specificity (CI 95%)**	**VPP (CI 95%)**	**VPN (CI 95%)**	**Sensitivity (CI 95%)**	**Specificity (CI 95%)**	**VPP (CI 95%)**	**VPN (CI 95%)**
Reader 1	**Set 1**	97.2% (90.3%-99.7%)	61.5% (31.6%-86.1%)	93.3% (82.1%-97.8%)	80% (44.4%-97.5%)	73.2% (57.1%-85.8%)	63.6% (47.8%-77.6%)	65.2% (49.8%-78.7%)	71.8% (55.1%-85%)
	Set 2	97.2% (90.3%-99.7%)	61.5% (31.6%-86.1%)	93.3% (82.1%-97.8%)	80% (44.4%-97.5%)	82.9% (67.9%-92.6%)	61.4% (45.5%-75.6%)	66.7% (52.1%-79.2%)	79.4% (62.1%-91.3%)
Reader 2	**Set 1**	88.9% (79.3%-95.1%)	23.1% (5%-53.8%)	86.5% (76.6%-93.3%)	27.3% (6%-61%)	73.2% (57.1%-85.8%)	70.5% (54.8%-83.2%)	69.8% (53.9%-82.8%)	73.8% (58%-86.1%)
	Set 2	88.9% (79.3%-95.1%)	23.1% (5%-53.8%)	86.5% (76.6%-93.3%)	27.3% (6%-61%)	92.7% (80.1%-98.5%)	47.8% (32.5%-63.3%)	62.3% (49%-74.4%)	87.5% (67.6%-97.3%)
Reader 3	**Set 1**	83.3% (72.7%-91.1%)	46.2% (19.2%-74.9%)	89.6% (79.7%-95.7%)	33.3% (13.3%-59%)	73.2% (57.1%-85.8%)	47.8% (32.5%-63.3%)	56.6% (42.3%-70.2%)	65.6% (46.8%-81.4%)
	Set 2	83.3% (72.7%-91.1%)	46.2% (19.2%-74.9%)	89.6% (79.7%-95.7%)	33.3% (13.3%-59%)	80.5% (65.1%-91.2%)	27.3% (15%-42.8%)	50.8% (38.1%-63.4%)	60% (36.1%-80.9%)

Table 3. ROC analysis results and pairwise comparison of the area under the curve (AUC) among the three readers with different experience.										
	Overall Cancer Detection		Clinically Significant Cancer Detection			Pairwise ROC Curve Comparison	Overall Cancer Detection		Clinically Significant Cancer Detection	
	AUC Set 1 (CI 95%)	AUC Set 2 (CI 95%)	AUC Set 1 (CI 95%)	AUC Set 2 (CI 95%)	p value		p value Set 1	p value Set 2	p value Set 1	p value Set 2
Reader 1	0.79 (0.69-0.87)	0.79 (0.69-0.87)	0.68 (0.57-0.78)	0.72 (0.61-0.81)	0.15	Reader 1 vs Reader 2	0.001	0.001	0.50	0.70
Reader 2	0.56 (0.45-0.67)	0.56 (0.45-0.67)	0.72 (0.61-0.81)	0.70 (0.59-0.80)	0.72	Reader 2 vs Reader 3	0.18	0.18	0.02	0.001
Reader 3	0.65 (0.54-0.75)	0.65 (0.54-0.75)	0.60 (0.49-0.71)	0.54 (0.43-0.65)	0.08	Reader 3 vs Reader 1	0.07	0.07	0.18	0.001

Figure 1. Flowchart of patients considered for inclusion, excluded patients and patients finally included in the study cohort.

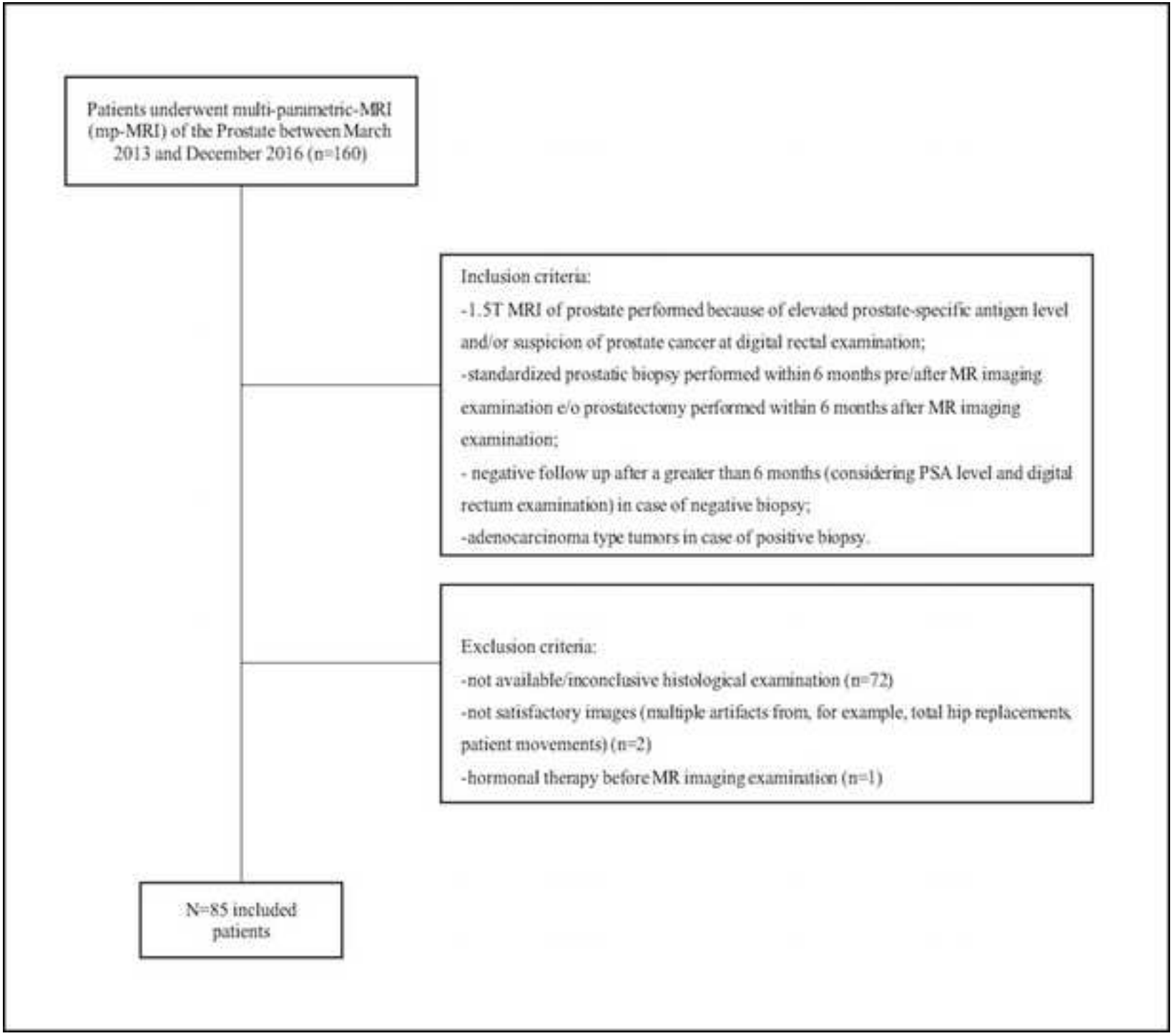
Figure 2. Comparison of ROC curves among the three readers regarding the overall cancer detection and clinically significant cancer detection using the two image sets (biparametric MRI and multiparametric MRI).

Figure 3. Example of correctly classified clinically significant tumor (Gleason Score 7 proven by radical prostatectomy). A focal hypointensity is visible in the right posterior peripheral zone on T2-weighted MRI. The lesion is markedly hyperintense on DWI and hypointense on ADC map and all the readers assigned a PI-RADS score of 4 for the Set 1 of MR images. DCE-MRI revealed an early and clear enhancement and did not changed the opinion of the readers in the Set 2 of MR images. Radical prostatectomy (RP) confirmed the presence of acinar adenocarcinoma (red arrow) with Gleason Score 7 (4+3) in the right posterior peripheral zone.

Figure 4. Example of uncorrectly classified non-clinically significant tumor (biopsy proven Gleason Score 6). A focal hypointensity is visible in the left posterior peripheral zone on T2-weighted MRI. The lesion is mild hyperintense on DWI and moderately hypointense on ADC map and the expert reader (Reader 1) assigned a PI-RADS score of 3 for the Set 1 of MR images. DCE-MRI revealed an early and clear enhancement by inducing the Reader 1 to increase the Gleason score to 4 for the Set 2 of MR images. The less expert readers (Reader 2 and 3) assigned a PI-RADS score of 4 for the Set 1 of images by interpreting the hyperintense signal on DWI and the hypointense signal on ADC map as focal and marked. They confirmed the PIRADS 4 score at DCE-MRI.

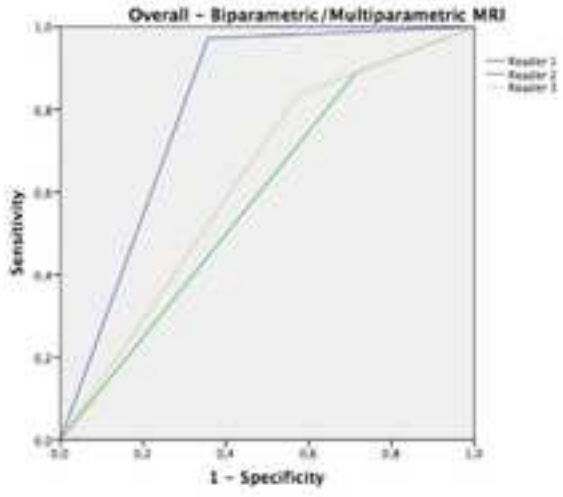
Figure(s)

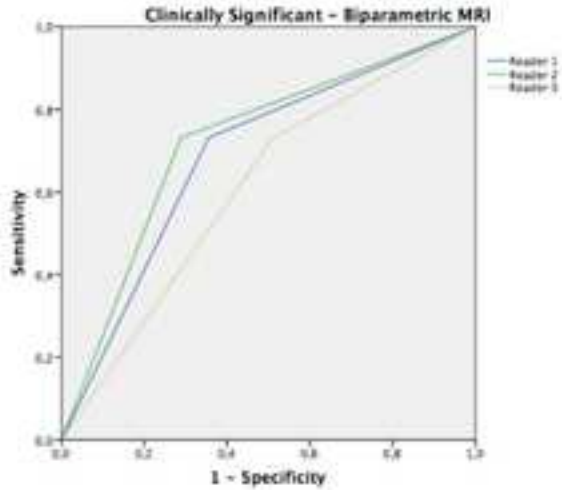
Click here to download high resolution image

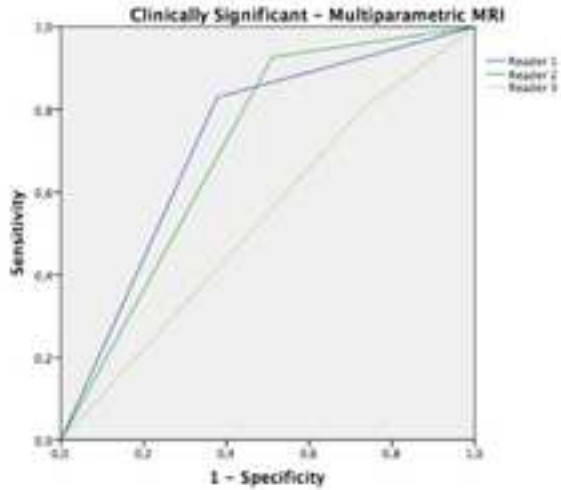


Figure(s)

Click here to download high resolution image

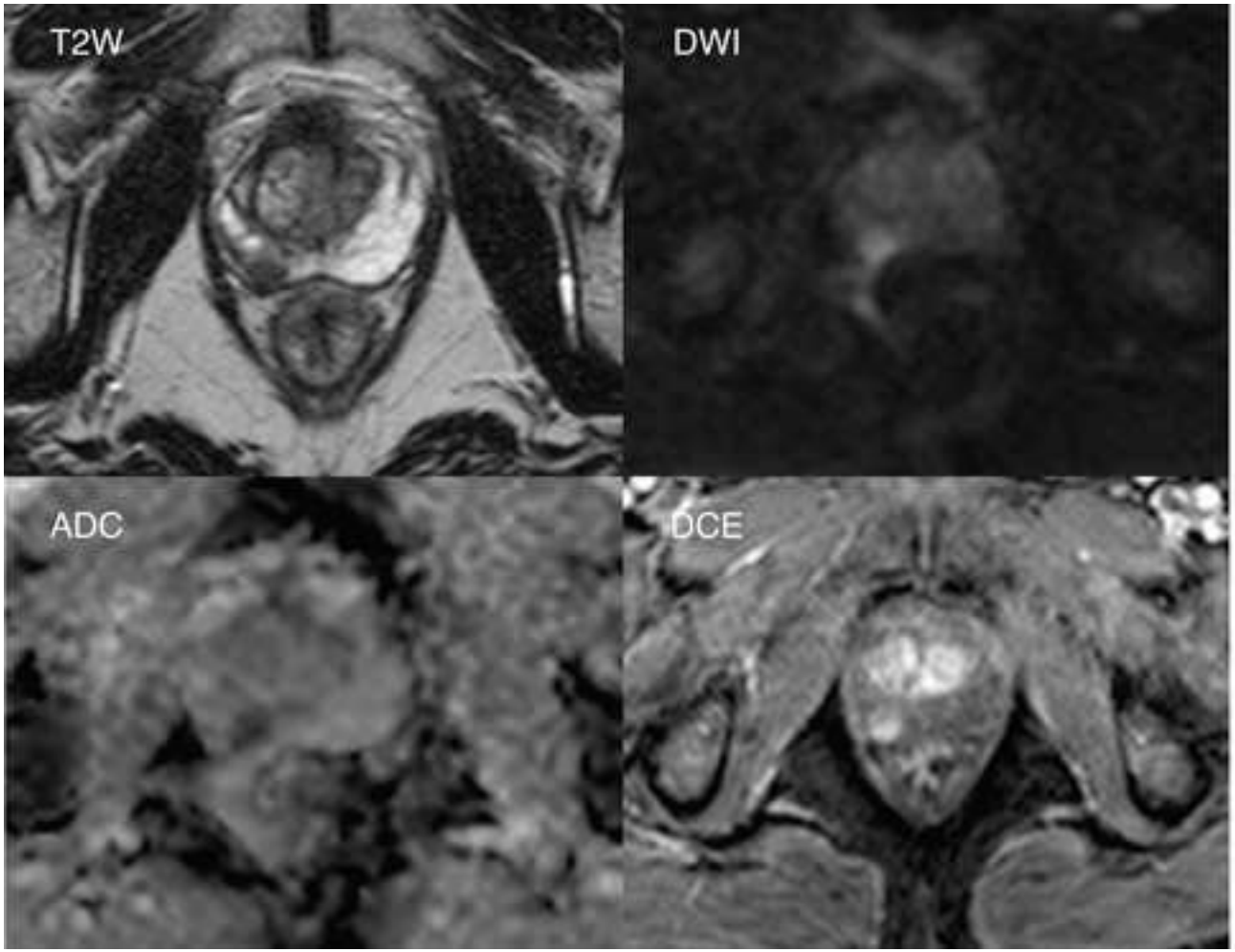


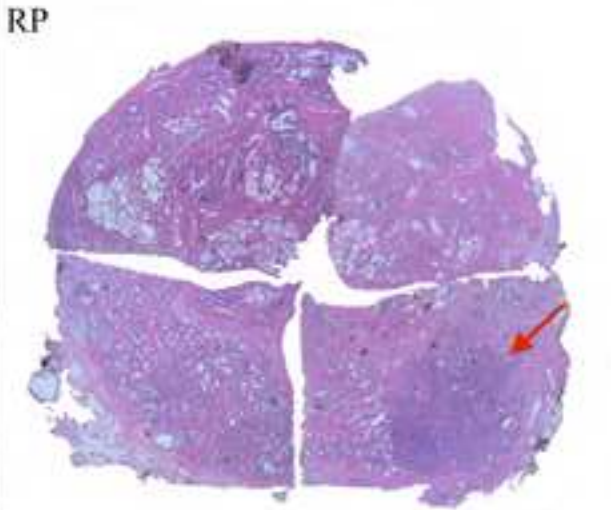




Figure(s)

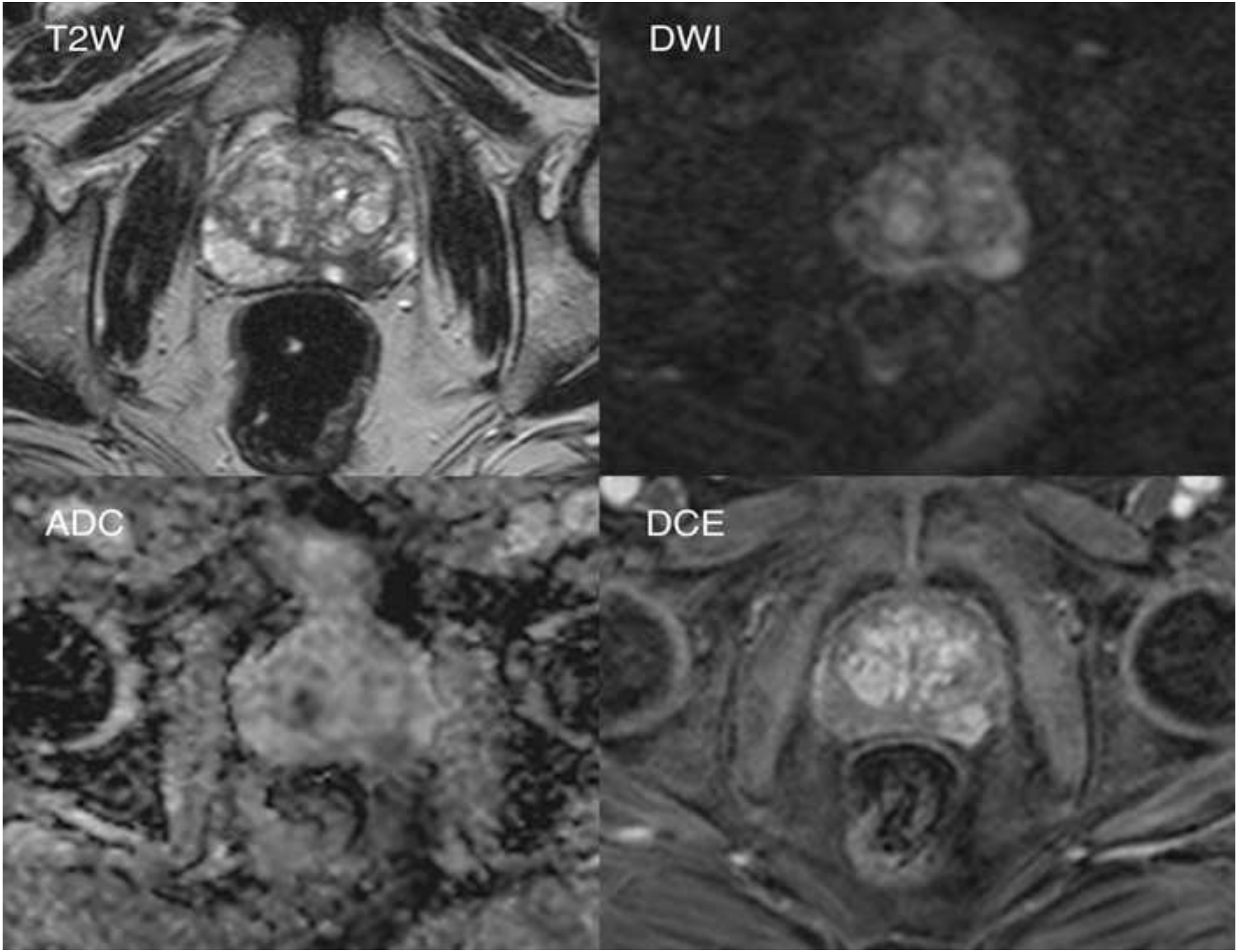
Click here to download high resolution image





Figure(s)

Click here to download high resolution image



Supplementary File(s)

Click here to download Supplementary File(s): Supplementary Material.docx