Variability of segmented prostate volume on MRI: impact on PSA density for prostate cancer diagnosis

RöFo - Fortschritte auf dem Gebiet der Röntgenstrahlen und der bildgebenden Verfahren

2025

10.1055/a-2624-7634

ISSN 1438-9029

Thieme

Variability of segmented prostate volume on MRI: impact on PSA density for prostate cancer diagnosis

Variabilität des segmentierten Prostatavolumens im MRT: Einfluss auf die PSA-Dichte für die Prostatakrebsdiagnose

Authors
Anna Colarieti[1], Alessandro Venturi[2], Massimo Cressoni[3], Matteo Interlenghi[2], Moreno Zanardo[4], Isabella Castiglioni[5], Francesco Sardanelli[6]

Affiliations
1 Department of Radiology, University of Eastern Piedmont Amedeo Avogadro School of Medicine, Novara, Italy
2 DeepTrace Technologies s.r.l., Milan, Italy, DeepTrace Technologies s.r.l., Milan, Italy
3 Department of Radiology, Policlinico San Donato, San Donato Milanese, Italy
4 Department of Radiology, San Donato Hospital Group, Milan, Italy
5 Department of Physics "G. Occhialini", Università degli Studi di Milano-Bicocca, Milan, Italy
6 Italian League for the Fight against Tumors (LILT) Milan Monza Brianza, Italian League for the Fight against Tumors (LILT) Milan Monza Brianza; Milan, Italy

Keywords
Digital Rectal Examination, Multiparametric Magnetic Resonance Imaging, prostate, Prostate-Specific Antigen, Prostatic Neoplasms

received 16.3.2025
accepted after revision 27.5.2025
published online 23.6.2025

Bibliography
Rofo
DOI 10.1055/a-2624-7634
ISSN 1438-9029

Georg Thieme Verlag KG, Oswald-Hesse-Straße 50, 70469 Stuttgart, Germany

Correspondence
Dr. Anna Colarieti
Department of Radiology, University of Eastern Piedmont Amedeo Avogadro School of Medicine, Via Solaroli 17, 28100 Novara, Italy
anna.colarieti@gmail.com

ABSTRACT

Purpose PSA density (PSAd), based on prostate volume (PV), is a decision-making parameter for prostate cancer (PCa) diagnosis and risk stratification. We assessed variability in prostate manual segmentation on MRI and its impact on PV and PSAd.

Materials and Methods We retrospectively analyzed 68 treatment-naïve patients, aged 66.2 ± 6.9 years, with increased PSA and/or positive digital endorectal examination who underwent MRI, with available biopsy/follow-up. Three radiologists (R1, R2, R3) manually segmented the gland on T2-weighted images slice-by-slice. Dice similarity coefficient (DSC), Welch's t-test, and 95% confidence intervals (CIs) were used.

Results Of 68 patients with a PSA of 7.59 ± 4.80 ng/mL, 38 had biopsy-confirmed PCa, and the remaining 30 were negative on biopsy/follow-up. The segmentation time per patient ranged from 4 to 7 min. Pairs R1-R2, R1-R3, and R2-R3 showed a different number of segmented slices ($p < 0.001$) and PV ($p < 0.001$). DSC for prostate gland segmentation ranged from 0.871 to 0.890. An outlier (prostatitis with PSA 35 ng/mL) was excluded from PSA/PSAd analysis. Based on segmentation by R1, the PSA was 7.37 ± 3.70 ng/mL and PSAd was 0.124 ± 0.070 ng/mL/mL in the 38 patients with PCa, while these values were 6.91 ± 2.79 ng/mL and 0.111 ± 0.062 ng/mL/mL, respectively, in the 29 patients without PCa. Using the threshold of ≥ 0.15 ng/mL/mL, variations in segmented PV impacted PSAd-based classification, resulting in 1 false negative for R1 and another false negative for R2 (false-negative rate for both 1/38, 2.63%, 95% CI 0.10–13.8%).

Conclusion Segmentation of PV is a time-intensive task. Inter-reader variability can impact PSAd-based diagnosis of PCa. Automated prostate segmentation methods are welcome.

Key Points

- Manual prostate segmentation is a time-consuming task performed in clinical practice
- Inter-reader variability of PV segmentation was low, with high DSC values
- Minor PV differences led to false-negative PSAd classification in 2.63% cases.

Citation Format

- Colarieti A, Venturi A, Cressoni M et al. Variability of segmented prostate volume on MRI: impact on PSA density for prostate cancer diagnosis. Rofo 2025; DOI 10.1055/a-2624-7634

ZUSAMMENFASSUNG

Zweck Die PSA-Dichte (PSAd), basierend auf dem Prostatavolumen (PV), ist ein entscheidungsrelevanter Parameter für die Diagnose von Prostatakrebs (PCa) und die Risikostratifikation. Wir untersuchten die Variabilität der manuellen Segmentierung der Prostata in der MRT und deren Einfluss auf PV und PSAd.

Material und Methoden Retrospektiv wurden 68 behandlungsnaive Patienten im Alter von 66,2 ± 6,9 Jahren mit erhöhtem PSA und/oder positivem digital-rektalem Befund, die sich einer MRT unterzogen hatten und bei denen Biopsie/Nachbeobachtung verfügbar war, analysiert. Drei Radiologen (R1, R2, R3) segmentierten die Drüse manuell scheibenweise auf T2-gewichteten Bildern. Dabei wurden der Dice-Ähnlichkeitskoeffizient (DSC), Welch's t-Test und 95 %-Konfidenzintervalle (CIs) verwendet.

Ergebnisse Von den 68 Patienten mit einem PSA von 7,59 ± 4,80 ng/mL hatten 38 einen bioptisch bestätigten PCa, während bei den übrigen 30 Patienten Biopsie/Nachbeobachtung negativ war. Die Segmentierungszeit pro Patient lag zwischen 4 und 7 Minuten. Die Paare R1–R2, R1–R3 und R2–R3 zeigten signifikante Unterschiede in der Anzahl segmentierter Schnitte ($p < 0{,}001$) sowie im PV ($p < 0{,}001$). Der DSC für die Segmentierung der Prostatadrüse lag zwischen 0,871 und 0,890. Ein Ausreißer (Prostatitis mit PSA 35 ng/mL) wurde aus der PSA/PSAd-Analyse ausgeschlossen. Basierend auf der Segmentierung durch R1 betrug bei den 38 PCa-Patienten das PSA 7,37 ± 3,70 ng/mL und die PSAd 0,124 ± 0,070 ng/mL/mL; bei den 29 Patienten ohne PCa lagen die Werte bei 6,91 ± 2,79 ng/mL bzw. 0,111 ± 0,062 ng/mL/mL. Bei Anwendung der Schwelle von ≥ 0,15 ng/mL/mL beeinflussten Variationen im segmentierten PV die PSAd-basierte Klassifikation, was zu einem falsch-negativen Befund bei R1 und zu einem weiteren falsch-negativen Befund bei R2 führte (falsch-negatives Rate für beide 1/38, 2,63 %, 95 %-CI 0,10–13,8 %).

Schlussfolgerung Die manuelle Segmentierung des PV ist sehr zeitintensiv. Die Inter-Reader-Variabilität kann die PSAd-basierte Diagnose des PCa beeinflussen. Methoden für eine automatisierte Segmentierung der Prostata sind daher willkommen.

Schlüsselpunkte

- Die manuelle Prostata-Segmentierung ist zeitaufwendig im klinischen Alltag.
- Geringe Variabilität zwischen Befundern bezüglich der PV-Segmentierung, mit hohen DSC-Werten.
- Kleine Volumenabweichungen führten zu 2,63 % falsch-negativen PSAd-Ergebnissen.

Abbreviations

DSC	Dice similarity coefficient
MRI	Magnetic resonance imaging
PI-RADS	Prostate Imaging Reporting and Data System
PCa	prostate cancer
PSA	prostate-specific antigen
PSAd	PSA density

Introduction

According to the 2020 data from the International Agency for Research on Cancer, prostate cancer (PCa) ranks as the second most diagnosed cancer globally, trailing only behind breast cancer, and is the fourth leading cause of cancer-related mortality, following lung, breast, and colorectal cancers [1, 2]. The diagnosis of PCa relies on a multifaceted approach including prostate-specific antigen (PSA) values, digital rectal examination, and imaging evaluation. Recently, there has been interest in multiparametric magnetic resonance imaging (MRI), due to its superior performance compared to transrectal ultrasound for PCa detection and the surveillance of high-risk patients [3].

PSA density (PSAd), computed as the ratio of PSA to prostate volume (PV), has emerged as a robust predictive factor for stratifying PCa risk [4], implying an obvious dependence of PSAd on PV. The approach to PV traditionally used for ultrasound [5, 6] relies on geometric approximated models based on simplified assumptions about the prostate's shape, considered similar to a regular ellipse-like structure, but this approach has low inter-reader reproducibility [7, 8, 9].

The current PI-RADS v2.1 guidelines mandate the inclusion of PV in MRI reports [10]. On MRI, PV can be determined through either manual or automated slice-by-slice segmentation or can be computed using the formula for a standard prolate ellipse [11, 12, 13]. Manual segmentation is recognized for providing an accurate volume estimation compared to pathological specimen volume [14, 15, 16] and is pivotal in PCa staging and treatment planning [17]. However, this task is time-intensive for expert radiologists due to the high morphological and volumetric variability of the gland in relation to age-related physiopathological changes as well as inflammatory conditions and malignancies [18, 19].

Few studies have examined the impact of MRI prostate segmentation on clinical management and treatment planning [20, 21, 22]. Prior experiences have mainly focused on intra- and inter-reader variability in total PV estimation [9, 23, 24, 25] and differentiation of gland zones [18] obtained by manual segmentation, with sample sizes ranging from 11 to 95 MRI examinations. However, to the best of our knowledge, this is the first study to compare inter-reader variability among three radiologists with different levels of prostate MRI experience.

Hence, the purpose of our study was to assess the inter-reader variability between radiologists with respect to PV manual segmentation on MRI and its impact on PSAd for PCa diagnosis.

▶ **Table 1** Technical parameters of axial T2-weighted sequences.

Parameter	Value
Repetition time (ms)	5540
Echo time (ms)	101
Flip angle (degrees)	150
Field of view (mm^2)	240 × 100
Matrix	256 × 110
Voxel size (mm^3)	1.1 × 0.9 × 3.0
Number of excitations	1

Materials and methods

The local ethics committee approved this retrospective study, and informed consent was waived due to the retrospective nature of the study and the irreversible data anonymization.

Study population

We included patients who underwent multiparametric MRI from January 2016 to December 2017: individuals aged 18 years or older who presented with a positive digital rectal examination and/or a PSA level exceeding 2.5 ng/mL. We excluded patients who had previously undergone interventions and/or treatments in the prostatic region (biopsy, surgery, embolization, or radiotherapy).

MRI protocol

All examinations were performed using a 1.5 T system (Magnetom Symphony Tim, Siemens Healthineers, Erlangen, Germany) equipped with an 8-channel phased-array coil and a single-channel endorectal coil. Prior to the examination, patients were required to undergo standard bowel preparation to optimize imaging quality. The multiparametric MRI protocol included T2-weighted sequences acquired in three spatial planes, axial diffusion-weighted imaging (DWI) sequences (*b* values of 50, 800, 1400 s/mm^2), dynamic contrast-enhanced study with intravenous gadolinium-based contrast agent using a three-dimensional T1-weighted axial sequence after contrast administration (0.1 mmol/kg of body weight) and a post-contrast T1-weighted axial sequence after the dynamic study. The technical parameters of the axial T2-weighted sequences on which the segmentation was performed are reported in ▶ **Table 1**.

Segmentation

Axial T2-weighted image datasets were irreversibly anonymized and imported into a research platform. All images were independently assessed by three radiologists (R1, R2, and R3 with 30 years, 6 years, and 1 year of urogenital imaging experience) to segment the prostate gland. After a consensus to establish the segmentation method, the radiologists segmented the PV. The platform presented the images as pairs of co-registered axial T2-weighted images (on the left) and apparent diffusion coefficient maps (on the right), allowing the reader to navigate up and down throughout the scanned volume. The prostate gland could be segmented on each T2-weighted image when visualized by the reader. The three readers agreed on the following steps: (i) identification of the prostate gland; (ii) determination of the apex (▶ **Fig. 1a**) and base (▶ **Fig. 1b**) of the gland; (iii) starting from the apex, navigating slice-by-slice down to the base to segment the gland in each slice where it was visible even though only for a minimal portion (▶ **Fig. 1c**); (iv) exclusion of the urethra and anterior fibromuscular tissue (▶ **Fig. 1d**). The PV values obtained by summing the slice-by-slice segmented areas multiplied by the slice thickness (3 mm) were used to calculate the PSAd as PSA/PV, expressed as ng/mL/mL.

Biopsy

In our study population, a single radiologist assessed the images in the clinical practice, including the characteristics of a suspicious prostate lesion including its score, size, and location based on PI-RADS v2.1. Patients with lesion scores of 3 or higher underwent either a targeted or systematic biopsy. A Gleason score of 7 or higher and the presence of extraprostatic invasion were used to define clinically significant or high-grade prostate cancer (PCa), specifically indicating a Gleason score > 3 and International Society of Urological Pathology grade > 3.

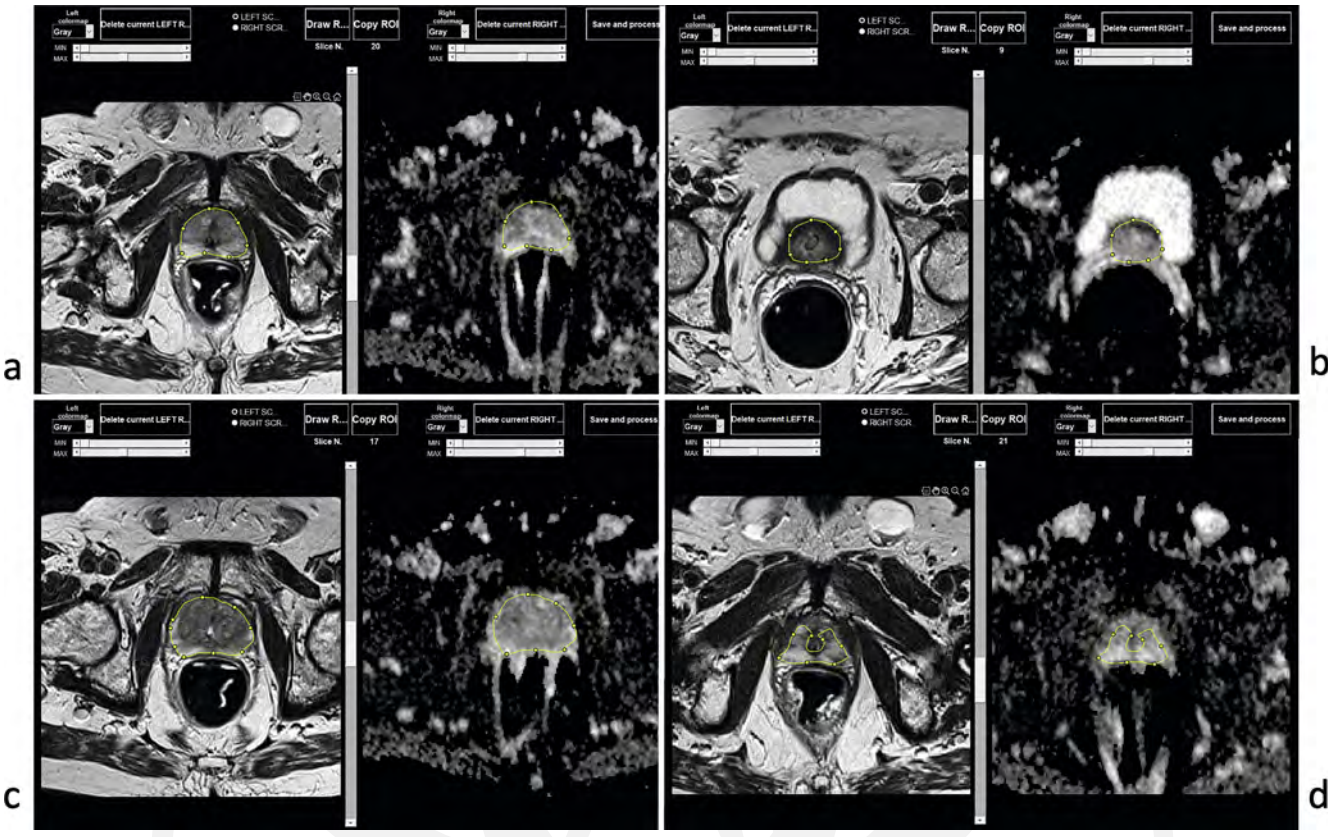
Statistical analysis

Normality of data distribution was verified through the Shapiro-Wilk test. Data are presented as mean ± standard deviation. The independent samples *t*-test was used to compare mean values between patients with negative and positive biopsy results.

The Dice similarity coefficient (DSC) was employed to evaluate the accuracy and consistency of prostate volume segmentations performed by different radiologists. The coefficient ranges from 0 to 1, where 0 indicates no overlap between two segmentations (complete disagreement) and 1 indicates perfect overlap (complete agreement). Welch's *t*-test was used to compare the means of two independent groups considering that the variances and sample sizes of the two groups were not equal; 95 % confidence intervals (CIs) were computed to quantify the uncertainty around the mean differences observed between the groups. We considered a DSC above 0.8 accurate in accordance with literature [18]. A threshold of ≥ 0.15 ng/mL/mL for PSAd positivity [26] was used. Statistical analyses were performed using Python software (numPy, pandas); *p*-values < 0.05 were considered significant [27].

Results

Patient demographic characteristics as well as the PSA values and Gleason score at biopsy are summarized in ▶ **Table 2**. Higher PSA values were observed in positive cases (7.37 ± 3.69 ng/mL) compared to the negative patient groups (6.91 ± 2.79 ng/mL), albeit without significant differences ($p = 0.556$). The PSAd based on PV measured by the segmentation of R1 (the most experienced operator) was 0.124 ± 0.070 for the positive group and 0.111 ± 0.062 ng/mL/mL for the negative one ($p = 0.423$).



▶ **Fig. 1** T2-weighted axial images and apparent diffusion coefficient (ADC) maps as shown by Trace4Research platform (DeepTrace Technologies S.R.L., Milan, Italy). **a** determination of the gland's apex; **b** determination of the gland's base; **c** segmentation of a slice located in the middle of the gland; **d** exclusion of the urethra and anterior fibromuscular tissue.

▶ **Table 2** Characteristics of the patient population.

Patient number	68
Age (years)	
▪ Mean ± standard deviation	66.2 ± 6.9
▪ Range	48–87
PSA (ng/mL)	
▪ Mean	7.59 ± 4.80
▪ Range	3.2–36
PI-RADS v2.1 (1–5)	
PI-RADS 1	1
PI-RADS 2	30
PI-RADS 3	7
PI-RADS 4	21
PI-RADS 5	9
Biopsy Gleason score	
Negative	30
▪ 6 (3 + 3)	11
▪ 7 (3 + 4 and 4 + 3)	17 (8; 9)
▪ 8 (3 + 5, 4 + 4, and 5 + 3)	8 (2; 4; 2)
▪ 9 (4 + 5)	2

Analyses involving pairs of readers (R1-R2, R1-R3, and R2-R3) showed differences in the number of segmented slices and in PV, as shown in ▶ **Table 3**. The DSCs for prostate gland segmentation were 0.884, 0.891, 0.867, respectively.

Regarding the PSAd based on the MRI PV, one patient (with a PSA = 35) was excluded from the subsequent analyses due to the final diagnosis of prostatitis as confirmed at follow-up with PSA < 2.5 ng/mL. Using the threshold of ≥ 0.15 ng/mL/mL, variations in segmented PV impacted on PSAd-based classification, resulting in two false negatives, one patient for R1 and another patient for R2 (false-negative rate 1/38, 2.63 %, 95 % CI 0.10–13.8 %) (▶ **Fig. 2**).

The segmentation time per patient was about 4 minutes for readers R1 and R2 and about 7 minutes for R3.

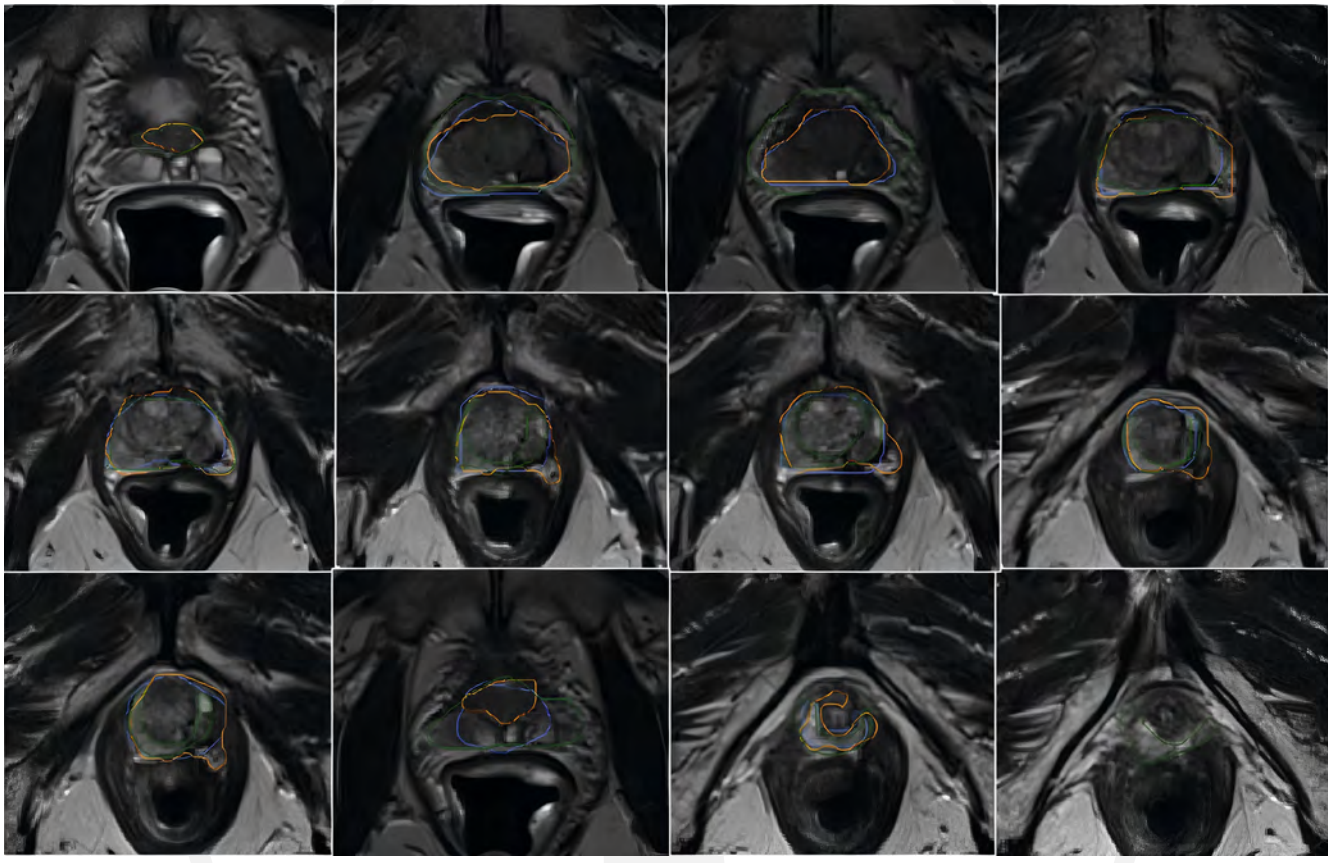
Discussion

In this study, we investigated the inter-reader variability for PV calculation on MRI among three radiologists, with the aim of minimizing PV variability as suggested by Moliere et al. [24]. We considered a subset of 68 patients, including 38 with biopsy-confirmed PCa and 30 with negative biopsy results or follow-up examinations. The assessment of inter-reader variability among radiologists (R1-R3, R1-R2, and R2-R3) revealed differences in segmented slices and PV (only for the pairs R1-R3 and R2-R3),

▶ **Table 3**

	R1	R2	R3	R1 vs. R2	R1 vs. R3	R2 vs. R3
Slices (*n*, mean ± SD)	18.3 ± 3.4	18.0 ± 3.9	19.1 ± 3.8	$p = 0.078$	$p < 0.001$	$p < 0.001$
Volumes (mL, mean ± SD)	70.6 ± 33.1	71.5 ± 33.7	67.8 ± 32.5	$p = 0.180$	$p < 0.001$	$p < 0.001$

SD = standard deviation



▶ **Fig. 2** Consecutive axial T2-weighted slices showing the segmentation performed by the 3 radiologists (R1 in blue, R2 in yellow, and R3 in green). The patient had a PSA of 3.40 ng/mL and a positive biopsy for csPCA. However, due to the calculated volume (R1: 22.5 mL, R2: 26.1 mL, R3: 19.8 mL), the PSA density values were 0.15, 0.13, and 0.17 ng/mL/mL, respectively, resulting in misclassification by R2.

even though the DSC was always equal to or higher than 0.860. The variation in the segmented PV impacted the PSAd-based classification for 2 out of 38 patients with PCa, misdiagnosed as PCa-negative when using PV segmented by R1 and R2 (false-negative rate of 2.63 %, 95 % CI 0.10–13.8 %). Variation in the segmented PV impacted the PSAd-based classification for 2 out of 38 patients with PCa. In both cases, the lesions were scored as PI-RADS 3, and the biopsy revealed Gleason scores of 3 + 4 and 4 + 3, respectively. These findings suggest that misclassification due to segmentation variability may delay diagnosis and treatment in patients with intermediate-risk prostate cancer, thus highlighting the real-world implications of segmentation variability.

Accurate measurement of PV is essential for managing conditions like benign prostatic hyperplasia, aiding in PCa diagnosis, and facilitating treatment planning [28]. PSAd plays a crucial role in identifying clinically significant prostate cancer (csPCa).

Prostate segmentation on MRI, typically performed on T2 W axial images by slice-by-slice contouring, is crucial in clinical applications, cancer staging, and treatment planning. However, this process is complex and time-consuming, including challenges associated with anatomical and physiological changes in the gland. Anomalies like abnormal hyperplasia growth, a third lobe protruding into the bladder, or the presence of tumors further complicate the segmentation task.

Numerous studies have explored the accuracy of PV estimation using ellipsoid formulas, comparing different measurement methods. Results generally indicate a high concordance between the PV calculated according to the ellipsoid formula and reference

standards, such as manual segmentation or prostatectomy specimens [29]. The study by Sosna et al. [11] included eleven patients imaged with a 3 T scan, employing T2-weighted imaging for PV calculation, who underwent radical prostatectomy. They compared PV estimations using ellipsoid and planimetric methods to the volumes of the pathologic specimens. Their results indicate that MRI-based PV estimations are highly accurate, especially when using the manual segmentation method, showing strong correlations with ex vivo measurements.

Ghafoor et al. [30] involved a retrospective analysis of 95 treatment-naive patients who underwent prostate MRI. In their study, PV was calculated using both ellipsoid and bullet formula methods, according to PI-RADS versions 2.0 and 2.1 guidelines. The results showed excellent inter-reader agreement for both formulas, with the ellipsoid method providing volume estimates closer to the manual reference standard. The research emphasized that the bullet formula tends to overestimate the real PV and highlighted that the ellipsoid measurement is more accurate and relevant for prostate volume estimation. However, there is a theoretical risk that using axial imaging to measure the anterior-posterior dimension of the prostate can lead to volume overestimation due to the gland's anatomical obliquity, which is referred to as the "salami effect". The latest PI-RADS update (v2.1) suggests using sagittal imaging to measure the anterior-posterior dimension to reduce this risk [31].

A recent study by Hamzaoui et al. [23] compared the PV of forty patients using the ellipsoid formula, biproximate ellipsoid formula, and manual segmentation. Both ellipsoid formulas consistently overestimated PV compared to manual segmentation. The manual segmentation, furthermore, had the highest inter-reader reproducibility. However, despite the accuracy of manual segmentation, variations in PV resulted in differing rates of patient classification discrepancies based on the clinical criterion of PSAd >0.15 ng/mL/mL, with manual segmentation showing the lowest rate of discrepancies at 5 %, similar to our false-negative rate of 2.6 % for two of the three readers.

We highlight our decision to conduct an interactive training meeting to align the segmentation methods, as suggested by Montagne et al [18]. In our opinion, this step is crucial given the gland's complexity, characterized by relevant variability in prostate volume, morphology, and different proportions between the transitional and peripheral zones. The presence or absence of benign and malignant lesions can also modify the gland's profile. Also, the exclusion of the anterior fibroglandular portion of the prostate was an object of agreement.

Despite minimal discrepancies in the DCS and total PV values among the three radiologists, a distinct segmentation approach resulted in a significant difference in the number of slices segmented (for the pairs R1-R3 and R2-R3). The two more experienced radiologists adopted a broader contouring strategy, encompassing fewer slices, while the less experienced radiologist adopted the opposite strategy. Indeed, R3 exhibited a more cautious contouring approach, with a mean segmentation time nearly double that of R1 and R2 (7 min. *versus* 4 min.). This result indicates the advantage of spending more time on segmentation in order to avoid misclassification of malignant cases and confirms that manual prostate segmentation on MRI remains a time-consuming task. Minor variations in PV segmentation can have a clinical impact.

We adopted a CsPCa threshold of 0.15 ng/mL/mL [26], which lead to a false-negative rate of 1/38 (2.6 %) for two of the three readers. However, this threshold was proposed by Di Donna et al. in 1993 [32], using PV estimation based on trans-ultrasound, when the use of MRI for prostate assessment was still in the early phase. According to Schoots et al. [33], PI-RADS 3 lesions with a low PSA density (<0.10 ng/mL/mL) have a low risk (4 %) of clinically significant disease, thus supporting imaging surveillance over biopsy. Conversely, their study revealed that 18 % of patients with PI-RADS 1 and 2 lesions but higher PSA density (>0.20 ng/mL/mL) show significant disease, suggesting a need for more intensive diagnostic strategies in these cases. For this reason, as suggested by Pellegrino et al. [34] there is a need to reconsider the PSAd threshold using the data derived from large MRI datasets.

Of note, for the analyses of PSA density, we excluded an outlier with a PSA of 35 ng/mL and mp-MRI and prostate biopsy negative for PCa. The patient had acute prostatitis, and the PSA value returned to normal (<2.5 ng/mL) three months after medical treatment.

Assessing inter-reader variability is crucial in the era of AI-based tools. Recent machine learning models have demonstrated excellent performance in prostate segmentation on MRI, with accuracy and reproducibility comparable to or exceeding that of human readers [35, 36]. However, challenges such as generalizability, translation from research software to CE-approved clinical tools, and integration into real clinical workflows still exist.

This study has several limitations. First, the relatively small sample size (only 68 patients in total, although 38 of these patients, i. e., 56 %, had clinically significant PCa) limits the statistical power and generalizability of the findings. Second, all patients underwent the same study protocol with the endorectal coil, which may introduce distortion regarding gland morphology and alter PV values due to compression. Finally, the limited sample size (68 patients in total, including 38 with csCPa) hindered precise estimation of the potential false-negative rate for PSAd, influenced by the variability in PV segmented by radiologists. We encountered a relatively broad 95 % confidence interval (0.10–13.8 %) around a point estimate of 2.6 %, suggesting an expected false-positive rate ranging from 1 in 7 to 1 in 10. Larger studies are certainly needed to address this relevant issue.

Conclusion

In conclusion, we showed that prostate manual segmentation is a time-consuming task requiring up to 7 minutes per patient and that minor variations in the obtained PV can result into a PSAd-based false-negative classification of over 2.6 % of malignant cases. These results warrant validation in larger prospective cohorts and should be interpreted within the future context of personalized PCa risk stratification.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

[1] Sung H, Ferlay J, Siegel RL et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 2021; 71: 209–249. doi:10.3322/caac.21660.

[2] Ferlay J, Colombet M, Soerjomataram I et al. Cancer statistics for the year 2020: An overview. Int J Cancer 2021. doi:10.1002/ijc.33588.

[3] de Rooij M, Israel B, Tummers M et al. ESUR/ESUI consensus statements on multi-parametric MRI for the detection of clinically significant prostate cancer: quality requirements for image acquisition, interpretation and radiologists' training. Eur Radiol 2020; 30: 5404–5416. doi:10.1007/s00330-020-06929-z.

[4] Sushentsev N, Abrego L, Colarieti A et al. Using a Recurrent Neural Network To Inform the Use of Prostate-specific Antigen (PSA) and PSA Density for Dynamic Monitoring of the Risk of Prostate Cancer Progression on Active Surveillance. Eur Urol Open Sci 2023; 52: 36–39. doi:10.1016/j.euros.2023.04.002.

[5] Mitterberger M, Horninger W, Aigner F et al. Ultrasound of the prostate. Cancer Imaging 2010; 10: 40–48. doi:10.1102/1470-7330.2010.0004.

[6] Huang Foen Chung JW, de Vries SH, Raaijmakers R et al. Prostate volume ultrasonography: the influence of transabdominal versus transrectal approach, device type and operator. Eur Urol 2004; 46: 352–356. doi:10.1016/j.eururo.2004.05.002.

[7] Nyholm T, Jonsson J, Soderstrom K et al. Variability in prostate and seminal vesicle delineations defined on magnetic resonance images, a multi-observer, -center and -sequence study. Radiat Oncol 2013; 8: 126. doi:10.1186/1748-717X-8-126.

[8] Aprikian S, Luz M, Brimo F et al. Improving ultrasound-based prostate volume estimation. BMC Urol 2019; 19: 68. doi:10.1186/s12894-019-0492-2.

[9] Becker AS, Chaitanya K, Schawkat K et al. Variability of manual segmentation of the prostate in axial T2-weighted MRI: A multi-reader study. Eur J Radiol 2019; 121: 108716. doi:10.1016/j.ejrad.2019.108716.

[10] Turkbey B, Rosenkrantz AB, Haider MA et al. Prostate Imaging Reporting and Data System Version 2.1: 2019 Update of Prostate Imaging Reporting and Data System Version 2. Eur Urol 2019; 76: 340–351. doi:10.1016/j.eururo.2019.02.033.

[11] Sosna J, Rofsky NM, Gaston SM et al. Determinations of prostate volume at 3-Tesla using an external phased array coil: comparison to pathologic specimens. Acad Radiol 2003; 10: 846–853. doi:10.1016/s1076-6332(03)00015-1.

[12] Lee DK, Sung DJ, Kim CS et al. Three-Dimensional Convolutional Neural Network for Prostate MRI Segmentation and Comparison of Prostate Volume Measurements by Use of Artificial Neural Network and Ellipsoid Formula. AJR Am J Roentgenol 2020; 214: 1229–1238. doi:10.2214/AJR.19.22254.

[13] Jeong CW, Park HK, Hong SK et al. Comparison of prostate volume measured by transrectal ultrasonography and MRI with the actual prostate volume measured after radical prostatectomy. Urol Int 2008; 81: 179–185. doi:10.1159/000144057.

[14] Morey RA, Petty CM, Xu Y et al. A comparison of automated segmentation and manual tracing for quantifying hippocampal and amygdala volumes. Neuroimage 2009; 45: 855–866. doi:10.1016/j.neuroimage.2008.12.033.

[15] Turkbey B, Fotin SV, Huang RJ et al. Fully automated prostate segmentation on MRI: comparison with manual segmentation methods and specimen volumes. AJR Am J Roentgenol 2013; 201: W720–729. doi:10.2214/AJR.12.9712.

[16] Smith WL, Lewis C, Bauman G et al. Prostate volume contouring: a 3D analysis of segmentation using 3DTRUS, CT, and MR. Int J Radiat Oncol Biol Phys 2007; 67: 1238–1247. doi:10.1016/j.ijrobp.2006.11.027.

[17] Bezinque A, Moriarity A, Farrell C et al. Determination of Prostate Volume: A Comparison of Contemporary Methods. Acad Radiol 2018; 25: 1582–1587. doi:10.1016/j.acra.2018.03.014.

[18] Montagne S, Hamzaoui D, Allera A et al. Challenge of prostate MRI segmentation on T2-weighted images: inter-observer variability and impact of prostate morphology. Insights Imaging 2021; 12: 71. doi:10.1186/s13244-021-01010-9.

[19] Chen ME, Troncoso P, Johnston D et al. Prostate cancer detection: relationship to prostate size. Urology 1999; 53: 764–768. doi:10.1016/s0090-4295(98)00574-3.

[20] Garvey B, Turkbey B, Truong H et al. Clinical value of prostate segmentation and volume determination on MRI in benign prostatic hyperplasia. Diagn Interv Radiol 2014; 20: 229–233. doi:10.5152/dir.2014.13322.

[21] Sanders JW, Mok H, Hanania AN et al. Computer-aided segmentation on MRI for prostate radiotherapy, Part I: Quantifying human interobserver variability of the prostate and organs at risk and its impact on radiation dosimetry. Radiother Oncol 2022; 169: 124–131. doi:10.1016/j.radonc.2021.12.011.

[22] Fernandes MC, Yildirim O, Woo S et al. The role of MRI in prostate cancer: current and future directions. MAGMA 2022; 35: 503–521. doi:10.1007/s10334-022-01006-6.

[23] Hamzaoui D, Montagne S, Granger B et al. Prostate volume prediction on MRI: tools, accuracy and variability. Eur Radiol 2022; 32: 4931–4941. doi:10.1007/s00330-022-08554-4.

[24] Moliere S, Hamzaoui D, Granger B et al. Reference standard for the evaluation of automatic segmentation algorithms: Quantification of inter observer variability of manual delineation of prostate contour on MRI. Diagn Interv Imaging 2024; 105: 65–73. doi:10.1016/j.diii.2023.08.001.

[25] Langkilde F, Masaba P, Edenbrandt L et al. Manual prostate MRI segmentation by readers with different experience: a study of the learning progress. Eur Radiol 2024. doi:10.1007/s00330-023-10515-4.

[26] Oishi M, Shin T, Ohe C et al. Which Patients with Negative Magnetic Resonance Imaging Can Safely Avoid Biopsy for Prostate Cancer? J Urol 2019; 201: 268–276. doi:10.1016/j.juro.2018.08.046.

[27] Di Leo G, Sardanelli F. Statistical significance: p value, 0.05 threshold, and applications to radiomics-reasons for a conservative approach. Eur Radiol Exp 2020; 4: 18. doi:10.1186/s41747-020-0145-y.

[28] Distler FA, Radtke JP, Bonekamp D et al. The Value of PSA Density in Combination with PI-RADS for the Accuracy of Prostate Cancer Prediction. J Urol 2017; 198: 575–582. doi:10.1016/j.juro.2017.03.130.

[29] Mazaheri Y, Goldman DA, Di Paolo PL et al. Comparison of prostate volume measured by endorectal coil MRI to prostate specimen volume and mass after radical prostatectomy. Acad Radiol 2015; 22: 556–562. doi:10.1016/j.acra.2015.01.003.

[30] Ghafoor S, Becker AS, Woo S et al. Comparison of PI-RADS Versions 2.0 and 2.1 for MRI-based Calculation of the Prostate Volume. Acad Radiol 2021; 28: 1548–1556. doi:10.1016/j.acra.2020.07.027.

[31] An JY, Fowler KJ. Accurate Prostate Volumes from Manual Calculations – A Comparison of PI-RADS v2 and v2.1 Measurement Techniques. Acad Radiol 2021; 28: 1557–1558. doi:10.1016/j.acra.2021.03.027.

[32] Di Donna A, Bazzocchi M, Guerra UP et al. The role of the absolute value and "density" of the prostate-specific antigen estimated echographically in the selection of patients to undergo a biopsy in suspected prostatic carcinoma. A comparison between PSA, palpation and echography in 95 patients undergoing echo-guided endorectal prostatic biopsy. Radiol Med 1993; 85: 84–89

[33] Schoots IG, Roobol MJ, Nieboer D et al. Magnetic resonance imaging-targeted biopsy may enhance the diagnostic accuracy of significant prostate cancer detection compared to standard transrectal ultrasound-guided biopsy: a systematic review and meta-analysis. Eur Urol 2015; 68: 438–450. doi:10.1016/j.eururo.2014.11.037.

[34] Pellegrino F, Tin AL, Martini A et al. Prostate-specific Antigen Density Cutoff of 0.15 ng/ml/cc to Propose Prostate Biopsies to Patients with Negative Magnetic Resonance Imaging: Efficient Threshold or Legacy of the Past? Eur Urol Focus 2023; 9: 291–297. doi:10.1016/j.euf.2022.10.002.

[35] Li H, Lee CH, Chia D et al. Machine Learning in Prostate MRI for Prostate Cancer: Current Status and Future Opportunities. Diagnostics (Basel) 2022; 12. doi:10.3390/diagnostics12020289.

[36] Correia ETO, Baydoun A, Li Q et al. Emerging and anticipated innovations in prostate cancer MRI and their impact on patient care. Abdom Radiol (NY) 2024; 49: 3696–3710. doi:10.1007/s00261-024-04423-4.