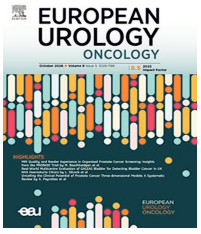
available at www.sciencedirect.com
journal homepage: euoncology.europeanurology.com



European Association of Urology

Original Article

MRI Quality and Reader Experience in Organized Prostate Cancer Screening: Insights from the PROBASE Trial

Matthias Boschheidgen [a], *Rouvier Al-Monajjed* [b,c], *Heinz-Peter Schlemmer* [d], *Agne Krilaviciute* [c], *David Bonekamp* [d], *Kathleen Herkommer* [e], *Matthias Jahnen* [e], *Jürgen E. Gschwend* [e], *Daniel Düx* [f], *Frank Wacker* [f], *Marcus R. Makowski* [g], *Andreas Sauter* [g], *Markus A. Kuczyk* [h], *Nina Harke* [h], *Jürgen Debus* [i], *Christoph Grott* [i], *Jale Lakes* [b], *Jan Philipp Radtke* [b,d], *Christian Arsov* [b,j], *Petra Seibold* [c], *Boris Hadaschik* [k], *Frederik Giesel* [l], *Glen Kristiansen* [m], *Axel Benner* [n], *Gerald Antoch* [a,o], *Nikolas Becker* [c], *Rudolf Kaaks* [p], *Peter Albers* [b,c], *Lars Schimmöller* [a,q,*]

[a] *University Dusseldorf, Medical Faculty, Department of Diagnostic and Interventional Radiology, Dusseldorf, Germany;* [b] *University Dusseldorf, Medical Faculty, Department of Urology, Dusseldorf, Germany;* [c] *Division of Personalized Early Detection of Prostate Cancer, German Cancer Research Center (DKFZ), Heidelberg, Germany;* [d] *Department of Radiology, German Cancer Research Center (DKFZ), Heidelberg, Germany;* [e] *Department of Urology, School of Medicine and Health, TUM university hospital rechts der Isar, Technical University of Munich, Munich, Germany;* [f] *Institute of Diagnostic and Interventional Radiology, Hannover Medical School, Hannover, Germany;* [g] *Department of Diagnostic and Interventional Radiology, School of Medicine and Health, TUM university hospital rechts der Isar, Technical University of Munich, Munich, Germany;* [h] *Department of Urology, Hannover Medical School, Hannover, Germany;* [i] *Department of Radiation Oncology, Heidelberg University Hospital, Ruprecht Karls University, Heidelberg, Germany;* [j] *Department of Urology and Paediatric Urology, Elisabeth-Krankenhaus Rheydt, Städtische Kliniken Mönchengladbach GmbH, Mönchengladbach, Germany;* [k] *Department of Urology, University of Duisburg-Essen and German Cancer Consortium (DKTK), University Hospital Essen, Essen, Germany;* [l] *University Dusseldorf, Medical Faculty, Department of Nuclear Medicine, Dusseldorf, Germany;* [m] *Institute of Pathology, University Hospital Bonn, Bonn, Germany;* [n] *Division of Biostatistics, German Cancer Research Center (DKFZ), Heidelberg, Germany;* [o] *Center for Integrated Oncology Aachen Bonn Cologne Dusseldorf (CIO ABCD), Germany;* [p] *Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany;* [q] *Department of Diagnostic, Interventional Radiology and Nuclear Medicine, Marien Hospital Herne, University Hospital of the Ruhr-University Bochum, Herne, Germany*

Article info

Article history:
Received 8 March 2026
Received in Revised form
31 March 2026
Accepted 5 May 2026
Available online 1 June 2026

Associate Editor:
Guillame Ploussard

Keywords:
Prostate cancer

Abstract

Background: The diagnostic performance of prostate magnetic resonance imaging (MRI) critically depends on image quality and reader expertise. The PROBASE trial is a prospective population- and prostate-specific antigen (PSA)–based prostate cancer (PCa) screening study enrolling men aged 45 yr.

Objective: In this predefined substudy, we evaluated the impact of MRI quality and expert reference reading on clinically significant PCa (csPCa; International Society of Urological Pathology grade ≥2) detection.

Design, setting, and participants: This analysis included 516 participants who underwent multiparametric MRI and combined MRI-targeted biopsy and systematic biopsy after screening with PSA ≥3 ng/ml. Local Prostate Imaging Reporting and Data System (PI-RADS) scores were compared with reference readings by experienced uroradiologists. MRI quality was assessed using Prostate Imaging Quality (version 1; PI-QUAL [v1]).

* Corresponding author. Department of Diagnostic and Interventional Radiology, University Dusseldorf, Medical Faculty, Moorenstr. 5 D-40225 Dusseldorf, Germany. Tel.: +49 - (0) 211 – 81 17754; Fax: +49 - (0) 211 – 81 16145.
E-mail address: lars.schimmoeller@med.uni-duesseldorf.de (L. Schimmöller).

https://doi.org/10.1016/j.euo.2026.05.003
2588-9311/© 2026 The Author(s). Published by Elsevier B.V. on behalf of European Association of Urology. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

MRI
PSA
MRI/US fusion biopsy
Risk-adapted screening

Outcome measurements and statistical analysis: Detection rates, positive predictive values (PPV), and negative predictive values (NPV) were calculated at an underlying biopsy threshold of PI-RADS score of ≥4. Image quality was compared between the study sites.
Results and limitations: Reference reading yielded higher NPV (92% vs 86%), PPV (57% vs 50%), and sensitivity (83% vs 69%) than local reading. Local reading missed substantially more csPCa in PI-RADS 1–2 (17.1% vs 2.9%) and resulted in more false-positive MRI findings in PI-RADS 4–5 (n = 76 vs n = 59) than reference reading. Lower MRI quality (PI-QUAL 1–3/5) was associated with reduced csPCa detection (46% vs 62%) and true negative rate (84% vs 95%). PI-QUAL differed between the study sites. PI-RADS classification of reference reading missed small or diffusely infiltrative csPCa, often in lower quality scans. Limitations include using local interpretations to define biopsy targets, limiting retrospective comparison, and possibly underestimating reference-reader performance.
Conclusions: Expert reference MRI reading substantially improves csPCa detection and can reduce unnecessary biopsies in a screening setting. High-quality imaging, standardized protocols, and experienced reading should be considered essential components of future MRI-based screening strategies.
Trial registration: The trial is registered with the ISRCTN (International Randomized Controlled Trial Number) registry, registration number ISRCTN37591328, and can be accessed at https://www.isrctn.com/ISRCTN37591328. The study protocol can be accessed at doi: 10.1016/j.eururo.2013.05.022.
© 2026 The Author(s). Published by Elsevier B.V. on behalf of European Association of Urology. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

ADVANCING PRACTICE

What does this study add?
This study provides robust evidence from a large, population-based screening cohort that both magnetic resonance imaging (MRI) image quality and reader expertise critically determine the diagnostic performance of prostate MRI examinations. Expert reference reading substantially improved the detection of clinically significant prostate cancer (PCa) while reducing false-positive findings and unnecessary biopsies compared with routine local interpretation. Lower image quality and heterogeneous reader experience were major contributors to missed or misclassified cancers in real-world screening conditions. These findings define essential prerequisites (standardized high-quality imaging and experienced interpretation) for safe and effective MRI-based PCa screening.

Clinical Relevance
In MRI-based prostate cancer screening, diagnostic accuracy is highly dependent on both image quality and reader expertise. Suboptimal image quality may lead to lower diagnostic performance and increased misclassification. In this population-based cohort of young men, expert MRI interpretation significantly improved detection of clinically significant prostate cancer while reducing false positives and unnecessary biopsies compared with routine local reading. These findings confirm that high-quality MRI acquisition, standardized protocols, and experienced reading (by centralized or perspectively with AI assistance) are essential prerequisites for safe and effective implementation of MRI-driven prostate cancer screening pathways. *Associate Editor: Guillame Ploussard.*

Patient Summary
In this study, we looked at how well prostate magnetic resonance imaging (MRI) works in a screening program for younger men. We found that MRI results are more accurate when images are of high quality and read by very experienced physicians. This helps detect important cancers earlier and can reduce unnecessary biopsies. However, some cancers can still be missed, especially when image quality is lower. Improving MRI quality and supporting doctors with tools such as artificial intelligence may make prostate cancer screening safer and more reliable in the future.

1. Introduction

The right strategy for a population-based screening program for prostate cancer (PCa) and, in particular, the role of magnetic resonance imaging (MRI) continues to be a controversial topic [1–5]. Several major studies included MRI in their design, with different results for an optimal concept. Some authors report the reduction in over-diagnosis, particularly when biopsy is limited to MRI-targeted cores and systematic biopsy is omitted [3,6–8]. However, systematic biopsy con-

tinues to identify PCa in a notable proportion of men with negative MRI findings, 19–28% of whom harbor clinically significant (cs) disease, defined as International Society of Urological Pathology (ISUP) grade ≥2 [9]. Especially in younger men, it seems to be difficult to leave out the systematic 12-core approach, as demonstrated earlier in our cohort [9]. The safety of deferring diagnosis in this subgroup remains uncertain [1,10]. It is still unclear whether the delayed detection of carcinomas of any ISUP group may preclude curative treatment, especially in a man aged 45 yr.

The critical role of reader experience in the interpretation of prostate MRI, particularly in the context of screening of younger men, has been mentioned before [11–13]. In our experience, expert reading resulted in the identification of 24% more csPCa cases when applying a threshold of Prostate Imaging Reporting and Data System (PI-RADS) ≥4 [1]. Besides this, image quality is a crucial factor in prostate MRI interpretation. PI-RADS version 2.1 (v2.1) sets stringent standards to ensure reliable and reproducible assessments [14]. With the Prostate Imaging Quality (PI-QUAL) score, it is possible to measure and compare image quality and set important standards that are crucial for quality assurance (currently updated to v2) [15]. When considering PCa screening at a population level, standardized acquisition protocols are necessary in multicenter screening settings [16–18]. The European quality assurance guidelines for breast cancer screening firmly establish independent double reading with consensus as the standard of care. This approach requires two readers to interpret mammograms independently while blinded to each other's decisions, followed by joint consensus resolution of any discordant findings. This method has demonstrated sustained value over many years, consistently increasing cancer detection rates, albeit with some increase in recall rates [19,20].

This study provides novel insights into MRI performance in a population-based screening cohort of young men, a setting that differs fundamentally from diagnostic cohorts. Specifically, it evaluates the combined impact of MRI reference reading by experienced uroradiologists and image quality on csPCa detection, identifies mechanisms of missed cancers, and quantifies the variability across centers in screening MRI examinations.

2. Materials and methods

2.1. PROBASE trial design

This analysis is based on the PROBASE screening trial. Between February 2014 and December 2019, a total of 46 495 eligible participants aged 45 yr were recruited from the general population and randomized into two screening groups in a 1:1 ratio: prostate-specific antigen (PSA) testing at the age of 45 years (immediate screening arm) versus delayed first PSA testing at the age of 50 years (deferred screening arm). Participants with confirmed elevated PSA levels of ≥3.0 ng/ml were recommended multiparametric MRI (mpMRI) of the prostate and prostate biopsy, independent of further risk classification. The PROBASE study was designed before the 2019 European Association of Urology guidelines recommended mpMRI for the primary diagnosis of PCa; therefore, biopsy in PROBASE is recommended to all participants with a confirmed PSA level of ≥3 ng/ml. Prostate biopsies were performed using targeted MRI/ultrasound (US) fusion-guided (TB) and systematic prostate biopsy (SB). All MRI examinations were performed at the four participating academic study centers (1: University Hospital Duesseldorf, 2: Hannover Medical School, 3: German Cancer Research Center/University Hospital Heidelberg, and 4: Technical University Munich). No external MRI studies were included in the analysis. mpMRI were read by local radiologists at each of the four study locations, who rated the images according to PI-RADS (for analysis, all classifications were documented along v2.1) and defined suspicious areas for the TB approach. In the case of PI-RADS categories 1 or 2, a representative area was defined as the target. The area should represent the image morphology at MRI as accurately as possible and be typical for the respective examination. In the second step, a central reference MRI reading was conducted retrospectively by two experienced radiologists in consensus not (primary) involved in the local image acquisition and interpretation or laggeing behind the initial reporting and blinded to histopathological results. Reference radiologists (L.S. and H.-P.S.) were board-certified radiologists from academic tertiary centers with dedicated expertise in uroradiology and received specialty training with Q2 (Qualification Level 2) certification of the Deutsche Röntgengesellschaft for reading prostate MRI. Reference readers had >10 yr of focused prostate MRI experience and had interpreted >10 000 prostate mpMRI examinations, with an annual case volume of >1000 studies. These expert radiologists evaluated the images according to PI-RADS (v2.1) and PI-QUAL v1. We systematically evaluated clinical, MRI (prostate volume, PSA density [PSAD], PI-RADS category, and PI-QUAL score), and biopsy data.

2.2. Study population

For this analysis, data from primary biopsies between February 2014 and September 2023 were extracted from the electronic central study database, Onkostar. Repeated biopsies from the same participants were excluded. All cases with external MRI and/or external biopsy were also excluded primarily because of the lack of traceability of the targeted cores and the lack of external biopsy and MRI reports. Thus, only participants with primary mpMRI and combined TB and SB at one of the four study sites were included.

2.3. Image acquisition and analysis

MRI parameters were chosen according to international recommendations. All scans were conducted on 3T MRI scanners. Examinations included T2WI turbo spin echo sequences in three planes, diffusion-weighted imaging (DWI) (b values of 0/50, 500, and 1000 with a calculated high b value ≥1400 s/mm^2), and dynamic contrast-enhanced imaging. Intravenous administration of a gadolinium-based contrast agent at a dosage of 0.1 mmol/kg body weight was conducted by default with a temporal resolution of ≤9 s. Interpretation of mpMRI was conducted locally by a board-certified radiologist who had access to the clinical details. There was no centralized review prior

to the biopsy. The likelihood of csPCa was evaluated using PI-RADS v2.1.

2.4. Biopsy

SB plus TB were performed using transrectal US and software fusion at all sites (transrectal or transperineal). For TB, a maximum of three lesions were defined. For SB, a scheme according to the size of the prostate was used, mostly of 12 independent systematic cores. TB was performed transrectally or transperineally. Biopsies were conducted using dedicated biopsy software (eg, Koelis, La Tranche, France; BiopSee, MedCom, Darmstadt, Germany; and UroNAV, Invivo, Gainesville, FL, USA). All biopsy procedures were performed by trained urologists at the study centers who were qualified for fusion biopsies, as described in a quality control standard procedure. Cancers with ISUP grade ≥2 were considered as csPCa. Histopathological analyses of the cores were conducted locally and were reviewed centrally by an experienced genitourinary pathologist with 20 yr of experience.

2.5. Outcome parameters, data, and statistical analysis

The primary outcome of this analysis was the diagnostic performance of prostate MRI for the detection of csPCa (ISUP grade ≥2), comparing local radiology interpretation with centralized expert reference reading. Patient, MRI, and biopsy data were analyzed descriptively. Median values with interquartile ranges (IQRs) were reported for screening PSA value, PSAD, PI-RADS, and prostate volume, and the median value together with IQR was reported for age at biopsy. Key performance metrics included sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) with confidence intervals, applying a biopsy threshold of PI-RADS ≥4. Secondary outcomes comprised the distribution of PI-RADS categories, interreader agreement between local and reference interpretation, and the frequency and characteristics of csPCa missed by either reading approach. Interreader agreement was calculated using weighted Cohen κ. Image quality was assessed using the PI-QUAL v1 score and analyzed across study sites and time. Associations between image quality and diagnostic performance metrics were evaluated to determine the impact of MRI acquisition quality on lesion classification and cancer detection. Subgroup analyses focused on csPCa detected in PI-RADS 1–3 categories and on discordant cases between local and reference reading. Clinical and imaging parameters, including prostate volume and PSAD, were compared between relevant subgroups. All analyses were performed on a per-patient basis. Statistical analyses were performed using IBM SPSS Statistics (v27; IBM Corp, Armonk, USA). Statistical significance was set at $p < 0.05$.

3. Results

3.1. Participant characteristics

From the overall cohort of 46 495 participants, 722 men (1.6%) have received a biopsy by September 2023, including men with repeated biopsy or PSA ≤3 ng/ml. A total of 592 of these participants underwent primary biopsy and had a screening PSA of ≥3.0 ng/ml; 67 participants were excluded because of external MRI, only TB or SB, or missing reference pathology. Of 525 participants in 9 MRI images were not available for reference radiology reading. Thus, the final study cohort consisted of 516 participants (Fig. 1). Baseline clinical characteristics are described in Table 1.

3.2. Local versus reference reading: PI-RADS distribution and csPCa detection

Table 2 illustrates the cross table for PI-RADS categories between local reading and reference reading. Interrater reliability was moderate, with a value of 0.46. PCa detection compared to PI-RADS for reference and local radiology is shown in Table 3. Taking a closer look at the PI-RADS stratification for reference reading, we had fewer cases of csPCa in PI-RADS categories 1–3 compared to those in local radiology results ($n = 25$ vs $n = 44$), and fewer negative cases in the PI-RADS categories 4–5 ($n = 59$ vs $n = 76$). However, for both readers, a high percentage of PI-RADS 3 cases were observed. Considering PI-RADS 3 as the cutoff for biopsy indication, there was no reduction in the number of negative biopsies compared to local radiology. For PI-RADS 3 cases at local reading, nonsignificant PCa ($n = 31$) and csPCa ($n = 31$) cases differed significantly in PSAD (0.13 ng/ml/ml vs 0.22 ng/ml/ml; $p \leq 0.01$), whereas prostate volume was similar for both groups (38 ml vs 40 ml).

Regarding the differences between local and reference radiology, we observed a higher NPV at PI-RADS categories 1–3 for the reference reading with an absolute difference of 6.3% and a higher PPV at PI-RADS categories 4–5 at a difference of 6.7% (Table 4, Fig. 2 and 3). Reference radiology missed 2.9% of csPCa within PI-RADS categories 1 and 2 compared to the 17.1% missed by local reading. Within the PI-RADS category 5, reference radiology had a very high rate of PCa in general (97.1%) and that of csPCa (94.3%). Furthermore, there were observed differences in detection rates between the study sites (Supplementary Tables 2, 3, 4, and 5).

3.3. Reference reading: csPCa in PI-RADS categories 1–3

Focusing on csPCa classified as PI-RADS categories 1–3 for reference reading, we observed 25 cases that exhibited cancers within ISUP grades 2–5 at a subsequent prostate biopsy. Six of the 25 participants revealed ISUP grade 3–5 cancers. Most of the MRI scans were of suboptimal image quality, with only six cases reaching a PI-QUAL v1 score of 4–5. In more than half of the cases, we had underlying diffuse changes at T2-weighted- and DWI-imaging within the peripheral zone (Table 5, Fig. 4 and 5). In 20% of the cases, cancerous areas could be visualized retrospectively when rereading the images, with a justified PI-RADS 4 score assigned.

3.4. Influence of image quality

Overall, in 382 of the 516 (74%) participants, image quality was optimal or good (PI-QUAL v1: 4–5), and in 134 of the 516 (26%) participants, it was acceptable to insufficient (PI-QUAL v1: 1–3). Although csPCa detection rates at reference reading did not differ significantly between PI-QUAL

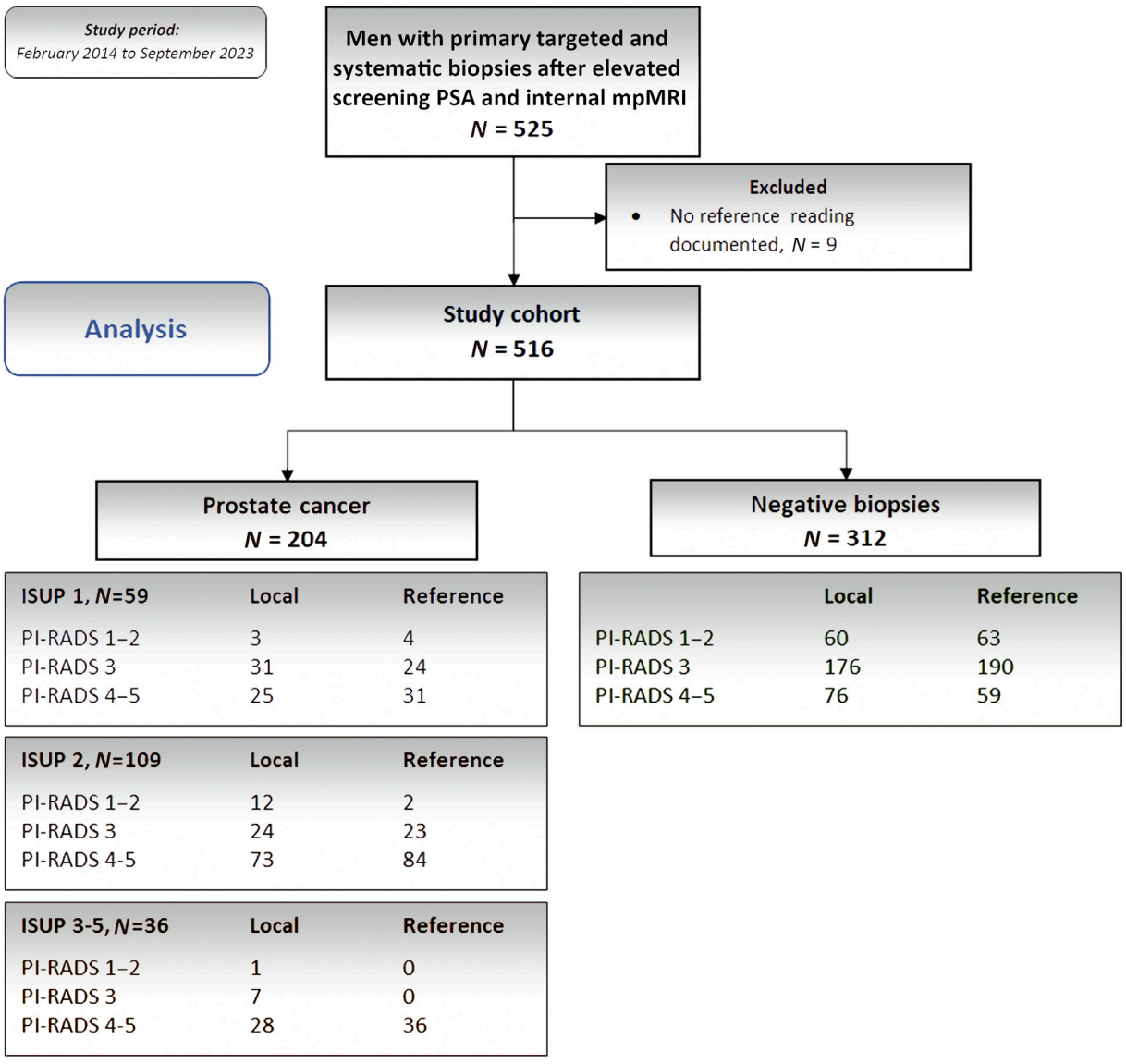


Fig. 1 – Flowchart of study design and participant selection. ISUP = International Society of Urological Pathology; mpMRI = multiparametric magnetic resonance imaging; PI-RADS = Prostate Imaging Reporting and Data System; PSA = prostate-specific antigen.

Table 1 – Baseline characteristics

Clinical characteristics	Total
Participants (*n*)	516
Age (yr), median (IQR)	50 (47–51)
PSA (ng/ml), median (IQR)	3.98 (3.45–5.08)
Prostate volume (ml), median (IQR)	36 (28–46)
PSAD (ng/ml/ml), median (IQR)	0.12 (0.09–0.16)

IQR = interquartile range; PSA = prostate-specific antigen; PSAD = prostate-specific antigen density.

1–3 and PI-QUAL 4–5 examinations (30.2% vs 27.0%), diagnostic performance metrics were inferior for lower quality MRI. CsPCa detection rate for PI-RADS 4–5 examinations was lower (45.5% vs 62.3%), and the true negative rate for PI-RADS ≤3 was reduced (83.6% vs 94.9%) in PI-QUAL 1–3 studies. Furthermore, the image quality showed differences between the study sites (Supplementary Table 6). However, an overall improvement in PI-QUAL scores was observed across the study period (during February 2014 to December 2020, median PI-QUAL score was 3 vs during January 2020 to September 2023, median PI-QUAL score was 4).

4. Discussion

In this large prospective screening cohort, we demonstrate a diagnostic advantage of expert MRI reference reading compared with local radiology reporting in the early detection of csPCa. Diagnostic performance could improve substantially with reference reading compared to local reading. A biopsy cutoff with PI-RADS ≥4 demonstrated relevant higher NPV, PPV, and sensitivity at reference reading. Moreover, it reduced both false positives and false negatives, leading to more precise classification of PI-RADS 4 and 5

Table 2 – Cross table of local and reference reading

PI-RADS		Local reading					Total
		1	2	3	4	5	
Reference reading	1	1	1	0	0	0	2
	2	1	32	25	8	1	67
	3	1	28	172	32	4	237
	4	2	10	40	117	6	175
	5	0	0	1	19	15	35
Total		5	71	238	176	26	516
Interrater reliability (Cohen κ): 0.46 (95% CI: 0.40–0.53)							

CI = confidence interval; PI-RADS = Prostate Imaging Reporting and Data System, version 2.1.

Table 3 – Cross table of PI-RADS classification and PCa detection for local and reference evaluation

Biopsies	Local PI-RADS classification				Total
	1 or 2	3	4	5	
Negative, *n* (%)	60 (78.9)	176 (74.0)	70 (39.8)	6 (23.1)	312
Positive, *n* (%)	16 (21.1)	62 (26.0)	106 (60.2)	20 (76.9)	204
ISUP GG 1, *n* (%)	3 (4.0)	31 (13.0)	22 (12.5)	3 (11.5)	59
ISUP GG 2–5, *n* (%)	13 (17.1)	31 (13.0)	84 (47.7)	17 (65.4)	145
Total	76	238	176	26	516
Biopsies	**Reference PI-RADS classification**				**Total**
	1 or 2	3	4	5	
Negative, *n* (%)	63 (91.3)	190 (80.2)	58 (33.1)	1 (2.9)	312
Positive, *n* (%)	6 (8.7)	47 (19.8)	117 (66.9)	34 (97.1)	204
ISUP GG 1, *n* (%)	4 (5.8)	24 (10.1)	30 (17.1)	1 (2.9)	59
ISUP GG 2–5, *n* (%)	2 (2.9)	23 (9.7)	87 (49.8)	33 (94.3)	145
Total	69	237	175	35	516

ISUP GG = International Society of Urological Pathology Grade Group; PCa = prostate cancer; PI-RADS = Prostate Imaging Reporting and Data System, version 2.1.

Table 4 – SEN, SPE, PPV, and NPV of local versus reference reading (biopsy cutoff PI-RADS ≥4 for csPCa)

csPCa	Local reading		Reference reading	
	Value (%)	95% CI (%)	Value (%)	95% CI (%)
SEN	69	62–77	83	76–89
SPE	73	67–77	76	71–80
PPV	50	45–55	57	52–62
NPV	86	82–89	92	89–94

CI = confidence interval; csPCa = clinically significant prostate cancer; PI-RADS = Prostate Imaging Reporting and Data System, version 2.1; PPV = positive predictive value; NPV = negative predictive value; SEN = sensitivity; SPE = specificity.

lesions and safer exclusion of csPCa in PI-RADS 1–3. Importantly, nearly half of all examinations were classified as PI-RADS 3, with no meaningful difference between local and reference interpretation, underscoring the intrinsic diagnostic ambiguity of this category. These findings provide evidence that expert reading not only enhances early csPCa detection but also has the potential to reduce unnecessary biopsies when applied upfront.

When interpreted by experienced radiologists, the detection rates for csPCa in our cohort approximate those reported in earlier nonscreening studies involving older populations [8,21,22]. This suggests that, despite the younger age and smaller gland volumes characteristic of PROBASE participants, expert MRI interpretation can achieve comparable diagnostic accuracy [23]. However, local readings performed under real-world conditions, without mandatory training or experience thresholds, yielded detection rates substantially below these benchmarks. This is consistent with prior work showing significant interreader variability and steep learning curves in prostate MRI interpretation [1]. Although interreader variability in prostate MRI has been widely described, our study extends this knowledge by demonstrating its clinical consequences in a population-based screening setting of young men. In contrast to diagnostic cohorts, lesion conspicuity is lower, and disease prevalence differs.

The lower performance of local reading likely reflects multiple design-related factors of the PROBASE protocol [24]. First, MRI examinations were conducted across multiple sites with differing scanners, evolving local protocols, and varying adherence to international performance standards. Although standard requirements were defined, optimization occurred over time, and MRI was not uniformly performed. This is emphasized by the technical improvement over time with improved image quality, as can be seen in the development of PI-QUAL scores. Second, local assess-

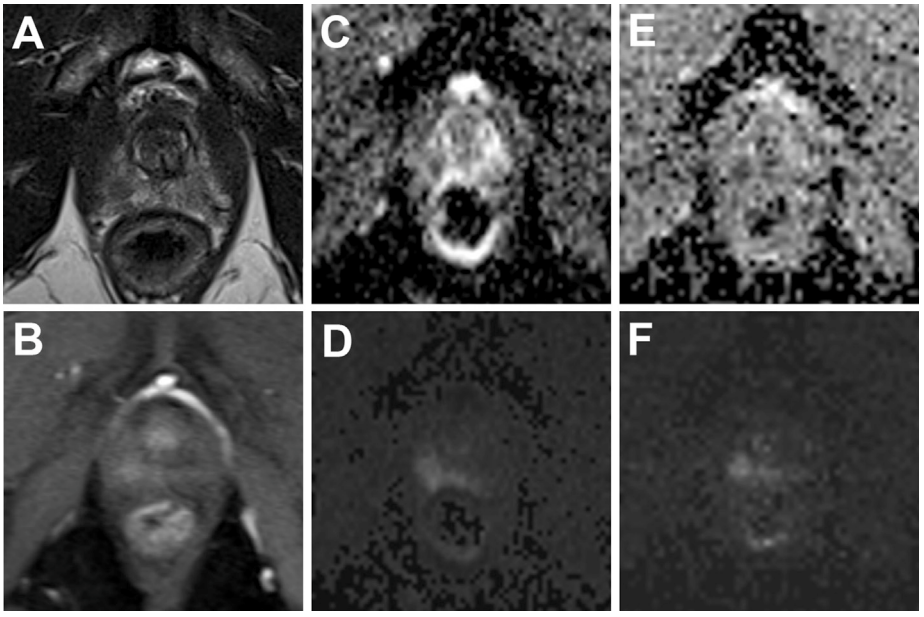


Fig. 2 – Study site 2 (PID 2224596), local PI-RADS 2, reference PI-RADS 4, PI-QUAL (v1) 4, ISUP 2, PSA 3.9 ng/ml, and PSAD 0.12 ng/ml/ml. (A) Axial T2-weighted image, (B) dynamic contrast-enhanced imaging, (C) ADC-Map of EPI-DWI, (D) high b-value image of EPI-DWI, (E) ADC-Map of RESOLVE-DWI, and (F) high b-value image of RESOLVE-DWI. ADC = apparent diffusion coefficient; EPI-DWI = echo-planar imaging diffusion-weighted imaging; ISUP = International Society of Urological Pathology; PID = Persistent Identifier; PI-QUAL (v1) = Prostate Imaging Quality score version 1; PI-RADS = Prostate Imaging Reporting and Data System; PSA = prostate-specific antigen; PSAD = prostate-specific antigen density; RESOLVE-DWI = readout segmentation of long variable echo-trains diffusion-weighted imaging.

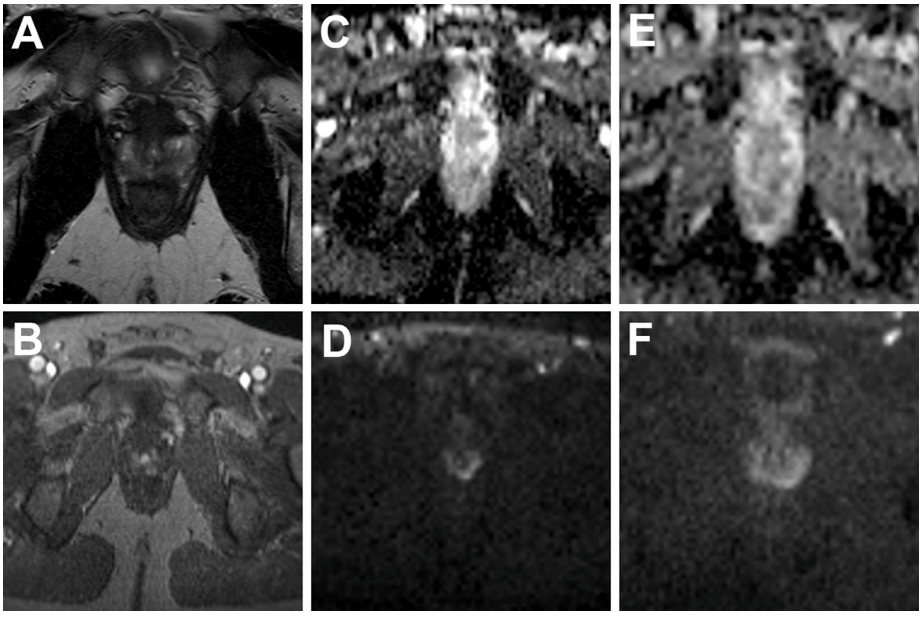


Fig. 3 – Study site 3 (PID 3347491), local PI-RADS 2, reference PI-RADS 4, PI-QUAL (v1) 2, ISUP 2, PSA 3.7 ng/ml, and PSAD 0.24 ng/ml/ml. (A) Axial T2-weighted image, dynamic contrast-enhanced imaging, (C) ADC-Map of EPI-DWI, (D) high b-value image of EPI-DWI, (E) ADC-Map of RESOLVE-DWI, and (F) high b-value image of RESOLVE-DWI. ADC = apparent diffusion coefficient; EPI-DWI = echo-planar imaging diffusion-weighted imaging; ISUP = International Society of Urological Pathology; PID = Persistent Identifier; PI-QUAL (v1) = Prostate Imaging Quality score version 1; PI-RADS = Prostate Imaging Reporting and Data System; PSA = prostate-specific antigen; PSAD = prostate-specific antigen density; RESOLVE-DWI = readout segmentation of long variable echo-trains diffusion-weighted imaging.

ment was performed by radiologists with heterogeneous experience levels, and no dedicated qualification (eg, Q2 certificate for prostate MRI reading) was required. Readers differed between study sites and even within sites, introducing substantial variability. Third, only lesions identified in the local read were biopsied; therefore, targets visible to the ref-

Table 5 – Cases missed by reference radiology with ISUP ≥2

Number	ISUP	PSA (ng/ml)	PSAD (ng/ml/ml)	Reference PI-RADS	Local PI-RADS	PI-QUAL v1	Retrospective comment
1	2	3.5	0.12	2	3	2	Very limited image quality, streaky peripheral changes, PI-RADS 3 confirmed
2	2	3.2	0.07	3	1	4	TZ/PZ interface, TB positive, PI-RADS 3 confirmed
3	2	4.6	0.11	3	2	3	Limited image quality, diffuse changes, TB positive, PI-RADS 3 confirmed
4	2	5.7	0.08	3	2	4	Atypical nodule, TB positive, PI-RADS 3 confirmed
5	2	5.6	0.28	3	2	2	Very limited image quality, older scanner, PI-RADS 3 confirmed
6	2	3.3	0.15	3	2	3	Limited image quality, small gland, small anterior apical, TB negative, PI-RADS 4 (on ADC/T2) feasible
7	2	3.8	0.09	3	2	3	Limited image quality, diffuse changes, TB negative, PI-RADS 3 confirmed
8	2	6.0	0.33	3	3	3	Limited image quality, diffuse changes, TB negative, PI-RADS 3 confirmed
9	2	3.0	0.08	3	3	4	Diffuse changes, TB negative, PI-RADS 3 confirmed
10	2	4.4	0.14	3	3	3	Limited image quality, diffuse changes, TB negative, PI-RADS 3 confirmed
11	2	3.3	0.09	3	3	4	Diffuse changes, questionable accentuation, TB positive, PI-RADS 4 (3 + 1) feasible
12	2	3.8	0.08	3	3	4	Diffuse changes, TB positive, PI-RADS 3 confirmed
13	2	3.4	0.07	3	3	3	Limited image quality, wedge shaped; PI-RADS 4 feasible
14	2	3.6	0.18	3	3	3	Limited image quality, diffuse changes, small gland, TB negative, PI-RADS 3 confirmed
15	2	3.1	0.08	3	3	3	Limited image quality, diffuse changes, TB negative, PI-RADS 3 confirmed
16	2	6.0	0.16	3	3	3	Limited image quality, diffuse changes, TB negative, PI-RADS 3 confirmed
17	2	3.5	0.18	3	3	3	Limited image quality, diffuse changes, questionable accentuation, TB positive, PI-RADS 4 (with rs-EPI-DWI)
18	2	3.7	0.12	3	4	3	Limited image quality, diffuse changes, TB negative, PI-RADS 3 confirmed
19	2	5.4	0.07	3	4	3	Limited image quality, subcapsular wedge-shaped; PI-RADS 4 (with rs-EPI-DWI)
20	3	4.2	0.14	2	3	3	Limited image quality, diffuse changes, TB negative, PI-RADS 3 confirmed
21	3	4.4	0.12	2	3	3	Limited image quality, diffuse changes, TB negative, PI-RADS 3 confirmed
22	4	4.5	0.11	3	2	4	Reference pathology positive, RPE ISUP 2, PI-RADS 3 confirmed
23	3	4.7	0.13	3	3	3	Limited image quality, RPE ISUP 2, PI-RADS 3 confirmed
24	3	7.2	0.40	3	3	3	Limited image quality, small gland, diffuse changes, PI-RADS 3 confirmed
25	5	4.2	0.13	3	3	3	Limited image quality, small gland, small subcapsular, TB negative, PI-RADS 4 (with rs-EPI-DWI)

ADC/T2 = apparent diffusion coefficient and T2-weighted imaging; DWI = diffusion weighted imaging; ISUP = International Society of Urological Pathology; PI-QUAL v1 = Prostate Imaging Quality score version 1; PI-RADS = Prostate Imaging Reporting and Data System; PSA = prostate-specific antigen; PSAD = prostate-specific antigen density; RPE = radical prostatectomy examination; rs-EPI = readout-segmented multishot echo-planar imaging; SB = systematic biopsy; TB = targeted biopsy; TZ/PZ = transition zone and peripheral zone.

erence reader, but not to the local reader, were not sampled, leading to an apparent inflation of missed csPCa in the local pathway. This could imply that PPV for reference reading and performance of TB might even be better.

Nonetheless, reference reading also failed to identify a subset of csPCa. These PCa were predominantly small, focal, or diffusely infiltrative, and many were also retrospectively not visible or only subtly detectable. This underlines a core limitation of prostate MRI, particularly in a young screening population with smaller glands and early-stage disease [12,25]. These characteristics may reduce lesion conspicuity, especially for low-visibility ISUP-2 PCa and those lacking a clear T2-weighted morphology. These findings also show the crucial importance of image quality [26,27]. Although PI-QUAL scores were acceptable in most examinations, our additional analysis suggests that image quality did differ systematically between the study sites and improved substantially over time. Examinations in which csPCa were missed tended to show lower PI-QUAL scores, indicating that even a slight to moderate reduction in image quality may particularly impact the detectability of subtle early-stage lesions. Standardized training, harmonized protocols,

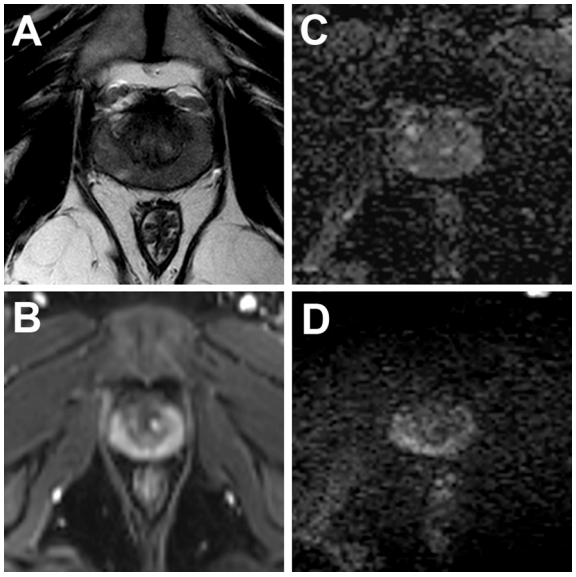


Fig. 4 – Study site 4 (PID 4988632), local PI-RADS 3, reference PI-RADS 3, PI-QUAL (v1) 3, ISUP 3, PSA 7.2 ng/ml, PSAD 0.40 ng/ml/ml. (A) Axial T2-weighted image, (B) dynamic contrast-enhanced imaging, (C) ADC-Map of EPI-DWI, and (D) high b-value image of EPI-DWI. ADC = apparent diffusion coefficient; EPI-DWI = echo-planar imaging diffusion-weighted imaging; ISUP = International Society of Urological Pathology; PID = Persistent Identifier; PI-QUAL (v1) = Prostate Imaging Quality score version 1; PI-RADS = Prostate Imaging Reporting and Data System; PSA = prostate-specific antigen; PSAD = prostate-specific antigen density; RESOLVE-DWI = readout segmentation of long variable echo-trains diffusion-weighted imaging.

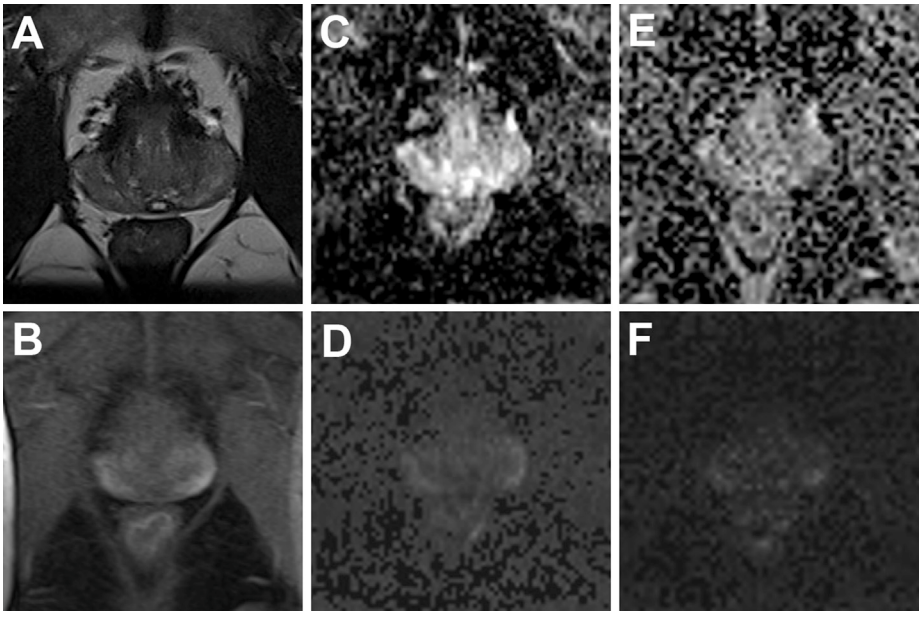


Fig. 5 – Study site 2 (PID 2968583), local PI-RADS 3, reference PI-RADS 3, PI-QUAL (v1) 3, ISUP 5, PSA 4.2 ng/ml, PSAD 0.13 ng/ml/ml. (A) Axial T2-weighted image, (B) dynamic contrast-enhanced imaging, (C) ADC-Map of EPI-DWI, (D) high b-value image of EPI-DWI, (E) ADC-Map of RESOLVE-DWI, and (F) a high b-value image of RESOLVE-DWI. ADC = apparent diffusion coefficient; EPI-DWI = echo-planar imaging diffusion-weighted imaging; ISUP = International Society of Urological Pathology; PID = Persistent Identifier; PI-QUAL (v1) = Prostate Imaging Quality score version 1; PI-RADS = Prostate Imaging Reporting and Data System; PSA = prostate-specific antigen; PSAD = prostate-specific antigen density; RESOLVE-DWI = readout segmentation of long variable echo-trains diffusion-weighted imaging.

and centralized quality assurance would help minimize such variability in future screening programs [16].

Given the high prevalence of PI-RADS 3 assessments (with predominant diffuse changes), refining this group seems crucial for a screening setting. PSAD provides a meaningful discriminator, with csPCa-positive PI-RADS 3 lesions showing significantly higher PSAD than the negative ones. Applying a PSAD cutoff of $\geq$0.15 ng/ml/ml would reduce

58 biopsies while still detecting eight additional csPCa cases. Nevertheless, 15 cases of csPCa would remain undetected, demonstrating that PSAD alone cannot fully resolve PI-RADS 3 uncertainty [28,29]. Prostate volume, in contrast, showed no significant difference, indicating limited utility for risk stratification in this age group.

In the local pathway, systematic biopsy added diagnostic value by compensating for missed MRI-visible lesions. It is also difficult to simply rescan PI-RADS 3 findings in this setting, as there is a relevant number of csPCa in this group, meaning that the cutoff for biopsy indication would probably have to be lowered. In contrast, the value of systematic sampling is likely reduced under expert MRI interpretation, as suggested by prior studies, demonstrating minimal incremental yield when MRI quality and expertise are optimized [9,25]. Therefore, in future screening models with expert/reference reading, systematic biopsies may be reserved for specific subgroups or eliminated entirely [3,7]. Furthermore, artificial intelligence (AI) may support scalable, quality-assured MRI-based PCa screening by standardizing image assessment, reducing interreader variability, and facilitating expert-level performance across heterogeneous clinical settings [30,31]. In this context, AI could serve as a triage or a second-reader system, particularly in settings where expert radiology resources are limited. Within the PROBASE study, there are already projects evaluating the utility of available AI tools for cancer diagnosis in this population.

Based on our findings, several key elements emerge for optimizing MRI within population-based screening:

- High and verifiable image quality using standardized, certified scanners and protocols.
- Reader expertise is certified by certain qualification levels and potentially double or reference reading to ensure consistency and to maximize csPCa detection.
- Restricting recommendations for immediate biopsy to PI-RADS 4–5, as these categories demonstrate a consistently high PPV under expert reference reading, especially when taking into consideration that missed csPCa were small, with a low percentage of Gleason 4 pattern, and are generally suitable for active surveillance strategies.
- Implementing a structured management pathway for PI-RADS 3, incorporating parameters such as PSAD, defining follow-up intervals, and, where appropriate, performing confirmatory or repeat imaging.

This analysis is constrained by the use of local interpretations to define biopsy targets, limiting retrospective comparison and possibly underestimating reference-reader performance. Local radiologists had varying levels of experience in interpreting prostate MRI scans, particularly in younger men, and some of them changed over the course of the study. Scanner heterogeneity and protocol evolution over time may also have influenced MRI quality. Because of inconsistent protocols at the start of the study and varying levels of qualification, there are discrepancies among the study centers. The learning curve over time may also lead to improved performance locally, which is not reflected here. Additionally, undetected lesions could not always be reliably characterized retrospectively. Furthermore, compared to previous studies, our sample size is limited; however, we are looking at a unique cohort that differs significantly from other screening cohorts in PCa screening.

5. Conclusions

This study demonstrates that expert reference MRI interpretation can markedly improve the early detection of csPCa while simultaneously showing the potential to reduce unnecessary biopsies. The reference read provided significantly higher PPV for PI-RADS 4 and 5 and higher NPV for PI-RADS 1–3, with a particularly large impact on improving lesion classification and avoiding unnecessary biopsies. These findings strongly support the integration of standardized high-quality imaging and specialized radiological expertise in any future MRI-based screening strategy and highlight the need for structured pathways tailored to younger populations and high proportions of PI-RADS 3 results.

***Author contributions*:** Lars Schimmöller had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Schimmöller, Boschheidgen, Al-Monajjed, Albers, Kaaks, Becker, Antoch.
Acquisition of data: Schimmöller, Boschheidgen, Schlemmer, Krilaviciute, Bonekamp, Debus, Grott, Gschwend, Jahnen, Kuczyk, Wacker, Düx, Harke, Makowski, Sauter, Kristiansen, Lakes, Herkommer, and the further PROBASE-Team.
Analysis and interpretation of data: Schimmöller, Boschheidgen, Al-Monajjed, Albers, Krilaviciute, Radtke, Hadaschik, Becker, Kaaks, Schlemmer, Jahnen, Bonekamp, Benner.
Drafting of the manuscript: Boschheidgen, Schimmöller, Al-Monajjed, Albers, Hadaschik, Antoch, Schlemmer.
Critical revision of the manuscript for important intellectual content: Albers, Hadaschik, Becker, Kaaks, Schlemmer, Radke, Antoch, Kristiansen, Seibold, Krilaviciute, Bonekamp, Düx, Arsov, Jahnen, Sauter, Giesel, Grott, Benner.
Statistical analysis: Boschheidgen, Schimmöller, Benner, Al-Monajjed, Krilaviciute, Becker, Kaaks, Albers.
Obtaining funding: Albers, Kaaks, Becker, Kuczyk, Debus, Gschwend, Seibold, Antoch, Hadaschik, Schimmöller, Kristiansen.
Administrative, technical, or material support: Albers, Kaaks, Becker, Krilaviciute, Seibold, Antoch, Schimmöller, Debus, Benner.
Supervision: Schimmöller, Albers, Antoch, Kaaks, Becker, Hadaschik.
Other (specify): None.

Financial disclosures: Lars Schimmöller certifies that all conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript (eg, employment/affiliation, grants or funding, consultancies, honoraria, stock ownership or options, expert testimony, royalties, or patents filed, received, or pending), are the following: None.

***Funding/Support and role of the sponsor*:** The PROBASE trial is fully funded by the German Cancer Aid (DKH; Stiftung Deutsche Krebshilfe). Open Access funding was enabled and organized by Projekt DEAL.

Acknowledgments: The authors want to thank the German Cancer Aid, namely, Gerd Nettekoven and Dr. Franz Kohlhuber, for their continuous and generous support of the trial. Open access funding was enabled and organized by Projekt DEAL. The authors want to thank Mrs. Kerstin Pieper, Mrs. Dorothee Zoller, and Dr. Sabine Behrens for their continuous and detailed support of the data management of the trial. Furthermore, the authors thank all documentary personnel currently working in the study sites and having contributed to the current data, namely, Nadja Bergen and Karin Thiess (Hannover), Iris Hillenbrand and Patrick Weis (Heidelberg), Simone Wachholz (Duesseldorf), and Helga Schulwitz (Munich), for their meticulous work on the raw data. In addition, the authors want to express their gratitude to documentary personnel who had worked for PROBASE in the past. In addition, the authors appreciate the support of the members of the scientific advisory board of the trial. Finally, the authors would like to thank Professor Monique Roobol, Erasmus University Rotterdam, for her critical remarks in reviewing the different steps of the development of the trial during the accrual phase.

Appendix A. Supplementary data

Supp. Tab. 1: Comparison of reference and local radiology for detection rates at different PI-RADS categories; Supp. Tab. 2: Comparison of PC detection and PI-RADS category in Study Site 1; Supp. Tab. 3: Comparison of PC detection and PI-RADS category in Study Site 2; Supp. Tab. 4: Comparison of PC detection and PI-RADS category in Study Site 3; Supp. Tab. 5: Comparison of PC detection and PI-RADS category in Study Site 4; Supp. Tab. 6: Difference between local and reference PI-RADS classification and PI-QUAL score (v1) distribution per Study site. Supplementary data to this article can be found online at https://doi.org/10.1016/j.euo.2026.05.003.

References

[1] Boschheidgen M, Albers P, Schlemmer HP, et al. Multiparametric magnetic resonance imaging in prostate cancer screening at the age of 45 years: results from the first screening round of the PROBASE trial. Eur Urol 2024;85:105–11.
[2] Auvinen A, Tammela TLJ, Mirtti T, et al. Prostate cancer screening with PSA, kallikrein panel, and MRI: the ProScreen randomized trial. JAMA 2024;331:1452–9.
[3] Hugosson J, Månsson M, Wallström J, et al. Prostate cancer screening with PSA and MRI followed by targeted biopsy only. N Engl J Med 2022;387:2126–37.
[4] Hugosson J, Roobol MJ, Månsson M, et al. A 16-yr Follow-up of the European Randomized study of Screening for Prostate Cancer. Eur Urol 2019;76:43–51.
[5] Hugosson J, Godtman RA, Wallstrom J, et al. Results after four years of screening for prostate cancer with PSA and MRI. N Engl J Med 2024;391:1083–95.
[6] Ahdoot M, Wilbur AR, Reese SE, et al. MRI-targeted, systematic, and combined biopsy for prostate cancer diagnosis. N Engl J Med 2020;382:917–28.
[7] Eklund M, Jäderling F, Discacciati A, et al. MRI-targeted or standard biopsy in prostate cancer screening. N Engl J Med 2021;385:908–20.
[8] Drost FH, Osses DF, Nieboer D, et al. Prostate MRI, with or without MRI-targeted biopsy, and systematic biopsy for detecting prostate cancer. Cochrane Database Syst Rev 2019;4:Cd012663.
[9] Al-Monajjed R, Boschheidgen M, Lakes J, et al. Prostate cancer detection in younger men: a comparative analysis of systematic and magnetic resonance imaging-targeted biopsy in the PROBASE trial. Eur Urol 2025;88:240–4.
[10] Al-Monajjed R, Schimmöller L, Radtke JP, et al. Characterisation of multiparametric magnetic resonance imaging of the prostate in younger men with normal prostate-specific antigen within the PROBASE study. Eur Urol Open Sci 2025;75:94–100.
[11] Kang HC, Jo N, Bamashmos AS, et al. Accuracy of prostate magnetic resonance imaging: reader experience matters. Eur Urol Open Sci 2021;27:53–60.
[12] Klingebiel M, Arsov C, Ullrich T, et al. Reasons for missing clinically significant prostate cancer by targeted magnetic resonance imaging/ultrasound fusion-guided biopsy. Eur J Radiol 2021;137:109587.
[13] Pausch AM, Ghafoor S, Kluckert J, Rupp NJ, Eberli D, Hötker AM. Risk factors for prostate cancer in men with false-negative mpMRI: a retrospective single center cohort study of image quality scores and clinical parameters. Eur J Radiol 2024;170:111227.
[14] 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;76:340–51.
[15] de Rooij M, Allen C, Twilt JJ, et al. PI-QUAL version 2: an update of a standardised scoring system for the assessment of image quality of prostate MRI. Eur Radiol 2024;34:7068–79.
[16] Schoots IG, Ahmed HU, Albers P, et al. Magnetic resonance imaging–based biopsy strategies in prostate cancer screening: a systematic review. Eur Urol 2025;88:247–60.
[17] Schoots IG, Padhani AR. A paradigm shift toward population-based MRI-targeted prostate cancer screening. JAMA Oncol 2025;11:828–30.
[18] Schoots IG, Haider MA, Punwani S, Padhani AR. MRI in prostate cancer screening: a review and recommendations, from the AJR special series on screening. AJR Am J Roentgenol 2025;225:e2432588.
[19] Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L. European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition–summary document. Ann Oncol 2008;19:614–22.
[20] Harvey SC, Geller B, Oppenheimer RG, Pinet M, Riddell L, Garra B. Increase in cancer detection and recall rates with independent double interpretation of screening mammography. AJR Am J Roentgenol 2003;180:1461–7.
[21] Rouvière O, Puech P, Renard-Penna R, et al. Use of prostate systematic and targeted biopsy on the basis of multiparametric MRI in biopsy-naive patients (MRI-FIRST): a prospective, multicentre, paired diagnostic study. Lancet Oncol 2019;20:100–9.
[22] Kasivisvanathan V, Rannikko AS, Borghi M, et al. MRI-targeted or standard biopsy for prostate-cancer diagnosis. N Engl J Med 2018;378:1767–77.
[23] Drost FJH, Osses D, Nieboer D, et al. Prostate magnetic resonance imaging, with or without magnetic resonance imaging-targeted biopsy, and systematic biopsy for detecting prostate cancer: a cochrane systematic review and meta-analysis. Eur Urol 2020;77:78–94.
[24] Arsov C, Becker N, Hadaschik BA, et al. Prospective randomized evaluation of risk-adapted prostate-specific antigen screening in young men: the PROBASE trial. Eur Urol 2013;64:873–5.
[25] Al-Monajjed R, Boschheidgen M, Lakes J, et al. Value of additional systematic cores during magnetic resonance imaging-guided targeted biopsy in prostate cancer screening for young men: results from the PROBASE trial. Eur Urol Oncol. In press. https://doi.org/10.1016/j.euo.2025.10.014.
[26] Thijssen LCP, Twilt JJ, Barrett T, et al. Quality of prostate MRI in early diagnosis-a national survey and reading evaluation. Insights Imaging 2025;16:82.
[27] Ponsiglione A, Cereser L, Spina E, et al. PI-QUAL version 2: a Multi-Reader reproducibility study on multiparametric MRI from a tertiary referral center. Eur J Radiol 2024;181:111716.
[28] Falagario UG, Martini A, Wajswol E, et al. Avoiding unnecessary magnetic resonance imaging (MRI) and biopsies: negative and positive predictive value of MRI according to prostate-specific antigen density, 4Kscore and risk calculators. Eur Urol Oncol 2020;3:700–4.
[29] Boschheidgen M, Schimmöller L, Doerfler S, et al. Single center analysis of an advisable control interval for follow-up of patients with PI-RADS category 3 in multiparametric MRI of the prostate. Sci Rep 2022;12:6746.
[30] Twilt JJ, Saha A, Bosma JS, et al. AI-assisted vs unassisted identification of prostate cancer in magnetic resonance images. JAMA Netw Open 2025;8:e2515672.
[31] Netzer N, Eith C, Bethge O, et al. Application of a validated prostate MRI deep learning system to independent same-vendor multi-institutional data: demonstration of transferability. Eur Radiol 2023;33:7463–76.