BJR, 2024, **97**, 1234–1242
https://doi.org/10.1093/bjr/tqae093
Advance access publication: 8 May 2024
Systematic Review

AI-accelerated prostate MRI: a systematic review

Ciaran Reinhardt , MB, BCh, BAO*[,1], Hayley Briody, MB, BCh, BAO[2], Peter J. MacMahon, MB, BCh, BAO[1]

[1]Department of Radiology, Mater Misericordiae Hospital, Dublin D07AX57, Ireland
[2]Department of Radiology, Beaumont Hospital, Dublin, D09V2N0, Ireland

*Corresponding author: Ciaran Reinhardt, MB, BCh, BAO, Department of Radiology, Mater Misericordiae Hospital, Dublin 7, D07AX57, Ireland (ciaranreinhardt@gmail.com)

Abstract

Background: Prostate cancer ranks among the most prevalent cancers affecting men globally. While conventional MRI serves as a diagnostic tool, its extended acquisition time, associated costs, and strain on healthcare systems, underscore the necessity for more efficient methods. The emergence of AI-acceleration in prostate MRI offers promise to mitigate these challenges.

Methods: A systematic review of studies looking at AI-accelerated prostate MRI was conducted, with a focus on acquisition time along with various qualitative and quantitative measurements.

Results: Two primary findings were observed. Firstly, all studies indicated that AI-acceleration in MRI achieved notable reductions in acquisition times without compromising image quality. This efficiency offers potential clinical advantages, including reduced scan durations, improved scheduling, diminished patient discomfort, and economic benefits. Secondly, AI demonstrated a beneficial effect in reducing or maintaining artefact levels in T2-weighted images despite this accelerated acquisition time. Inconsistent results were found in all other domains, which were likely influenced by factors such as heterogeneity in methodologies, variability in AI models, and diverse radiologist profiles. These variances underscore the need for larger, more robust studies, standardization, and diverse training datasets for AI models.

Conclusion: The integration of AI-acceleration in prostate MRI thus far shows some promising results for efficient and enhanced scanning. These advancements may fill current gaps in early detection and prognosis. However, careful navigation and collaborative efforts are essential to overcome challenges and maximize the potential of this innovative and evolving field

Advances in knowledge: This article reveals overall significant reductions in acquisition time without compromised image quality in AI-accelerated prostate MRI, highlighting potential clinical and diagnostic advantages.

Keywords: artificial intelligence; MRI; prostate cancer; AI-accelerated imaging.

Introduction

Prostate cancer is the second most common, and fourth leading cause of cancer-related mortality in men worldwide.[1] In the USA alone, incidences of prostate cancer rose by 3% annually between 2014 and 2020.[2] Given the impact on outcome and longer-term prognosis, early detection in prostate cancer is paramount, however no acceptable form of screening exists.

Previously, MRI was used primarily in patients who had biopsy-proven prostate cancer, however, MRI is now used for detection, localization, staging, treatment guidance, and surveillance.[3,4] Acquisition of the multiparametric (mpMRI) protocol, as recommended by Prostate Imaging Reporting and Data System (PIRADS), conventionally takes between 30 and 45 min to complete.[5,6] Of the four main imaging techniques in mpMRI, T2WI and DWI play a central role in lesion detection and characterization.[7,8] T2WI accounts for the bulk of this time, constituting approximately 60% of the total overall acquisition time. Although this allows detailed imaging, it poses challenges such as patient discomfort, anxiety, and motion-induced artefacts that reduce image quality.[9] Longer scans also limit daily throughput, affecting efficiency and cost.

Clinically, efforts to reduce artefacts and scan time, and improve contrast and spatial resolution, have manifold repercussions on patient care. Shorter scans may enhance patient comfort and allay anxiety, while also translating to increased departmental throughput, thus potentially addressing the perennial challenges of protracted waiting lists.[10,11] Higher image quality ensures accurate diagnosis and better risk stratification, along with heightened accuracy in tumour diagnoses and characterization. Cumulatively, enhanced MRI techniques can not only bolster diagnostic certainty but also translate to tangible benefits to patients including improved treatment outcomes and enhanced patient experience.[8]

In earlier attempts to address challenges in MRI acquisition, parallel imaging (PI) techniques emerged as the preferred technique used to reduce scan time while aiming to preserve image fidelity. These techniques, though groundbreaking in their time, were not without limitations.[12] As the quest for higher acceleration grew, performance and fidelity issues became more pronounced. Newer techniques began to surface, offering alternative solutions to reduce scan time. Noteworthy among them was compressed sensing (CS). The biggest drawbacks of CS in widespread clinical adaptation, however, are that it is computationally demanding in terms of processing, it generates its own artefacts from inaccurate sparse modelling assumptions, its parameter settings need to change depending on the application, and its reconstruction timings have proven clinically infeasible.[13,14]

Received: 15 November 2023; **Revised:** 5 April 2024; **Accepted:** 4 May 2024

With emerging AI advancements, radiology stands to benefit immensely. AI's core strength is processing large data sets, recognizing intricate patterns, and refining algorithms. Tailored to MRI, AI promises faster acquisition protocols, better reconstruction methods, and some automated interpretation facets. Crucially, AI looks to be able to speed up MRI without sacrificing image quality.[15]

In acquisition, AI can interpolate missing data, speeding up data collection. With reconstruction, deep learning (DL) often outperforms traditional methods.[16] Preliminary research hints at AI's ability to distinguish benign from malignant, along with predicting tumour aggressiveness and forecasting treatment.[17,18] As AI models continue to evolve, it is plausible they will offer radiologists enhanced diagnostic accuracy and reduced inter-observer variability.

This review will assess, evaluate, and synthesize the extant literature from studies of AI-accelerated prostate MRI, and will evaluate the following key dimensions:

1) Reduction in scan duration: Quantifying the time savings achieved through AI-accelerated prostate MRI which will consider, in particular, whether the reduction in time maintains diagnostic integrity.
2) Qualitative measurement: Exploration of AI's impact on image clarity, resolution, and contrast, including consideration of artefact reduction, noise reduction, and image sharpness and/or diagnostic confidence.
3) Quantitative measurement: Objective determination of whether AI-augmented prostate MRI results in more, less, or objectively similar signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR).

Methods

This systematic review is reported in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 checklist,[19] with a completed checklist provided in Supplementary Appendix F. This research did not require ethical approval. Papers were also assessed using the Checklist for Artificial Intelligence in Medical Imaging (CLAIM).[20]

Search strategy

Independent electronic searches were conducted on PubMed, Embase, and Scopus for trials or studies on AI-accelerated MRI prostate imaging from January 1, 2013 to August 10, 2023. Only English publications or those with English translations were considered. Specific search terms for each database included "Artificial Intelligence", "Machine Learning", "Magnetic Resonance Imaging", "Prostate", and "Acceleration", along with more granular terms such as "Convoluted Neural Network" and "Deep Learning". Full search terms are detailed in the Supplementary material.

Study selection

Articles from searches were added to Covidence,[21] and auto-de-duplicated. Two reviewers then manually screened abstracts, resolving conflicts through mutual discussion. Relevant full texts were subsequently analysed.

Selection criteria

Inclusion criteria

Studies from 2013 to 2023 using AI for MRI acceleration with primary outcomes of acquisition time and image quality. Secondary outcomes encompassed diagnostic confidence, lesion visibility, and reader consistency.

Exclusion criteria

Excluded were case reports, letters, editorials, opinions, commentaries, reviews, and non-prostate MRI studies. Reasons for exclusion during full-text review included wrong interventions such as 3D, combination interventions such as compressed sense and AI, and reconstruction techniques.

Data extraction

Two reviewers independently extracted technical attributes, and quantitative and qualitative indices from each study.

Assessment of risk of bias

The QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies-2) tool,[22] while specifically designed to evaluate the quality and potential biases in diagnostic accuracy studies, has been customized to suit the needs of this review. This involved incorporating specific criteria to evaluate the reliability, accuracy, and potential biases inherent in AI technologies, such as assessing the quality of data used to train the AI models, and their ability to generalize across different patient populations. Additionally, the tool was adapted to ensure that AI diagnostics were compared against universally accepted benchmarks, like those recommended by PIRADS for mpMRI acquisition.

Data analysis

Due to varied algorithms and validation strategies among studies, a narrative synthesis was preferred over a meta-analysis to prevent bias.

Results

Study selection

Figure 1 details the study selection process. One hundred and fifty-two papers were identified and imported into Covidence. Sixty-five duplicate papers removed following manual verification. Eighty-seven papers were screened, of which 75 were excluded. Twelve papers underwent full-text screening, with four papers excluded due to incorrect interventions. Eight studies were included and underwent CLAIM, QUADAS-2 assessment, data extraction, and narrative synthesis.

Quality review

The CLAIM checklist assessed the quality of eight studies,[23-30] with all meeting over 50% of the criteria, highlighting quality in AI research. The varied applicability in AI imaging may explain these differences, with some criteria more suited for diagnostics. While all studies demonstrated substantial conformity to the checklist, it is crucial to recognize these nuances in AI research domains. Full checklist details are in the Supplementary material.

Risk of bias assessment

All studies included show a low risk of bias on a modified QUADAS-2 assessment. Studies were rated on a 3-point scale and were classified as either low (+), high (−), or unclear (?). Details are shown in Table 1. In the Patient Selection

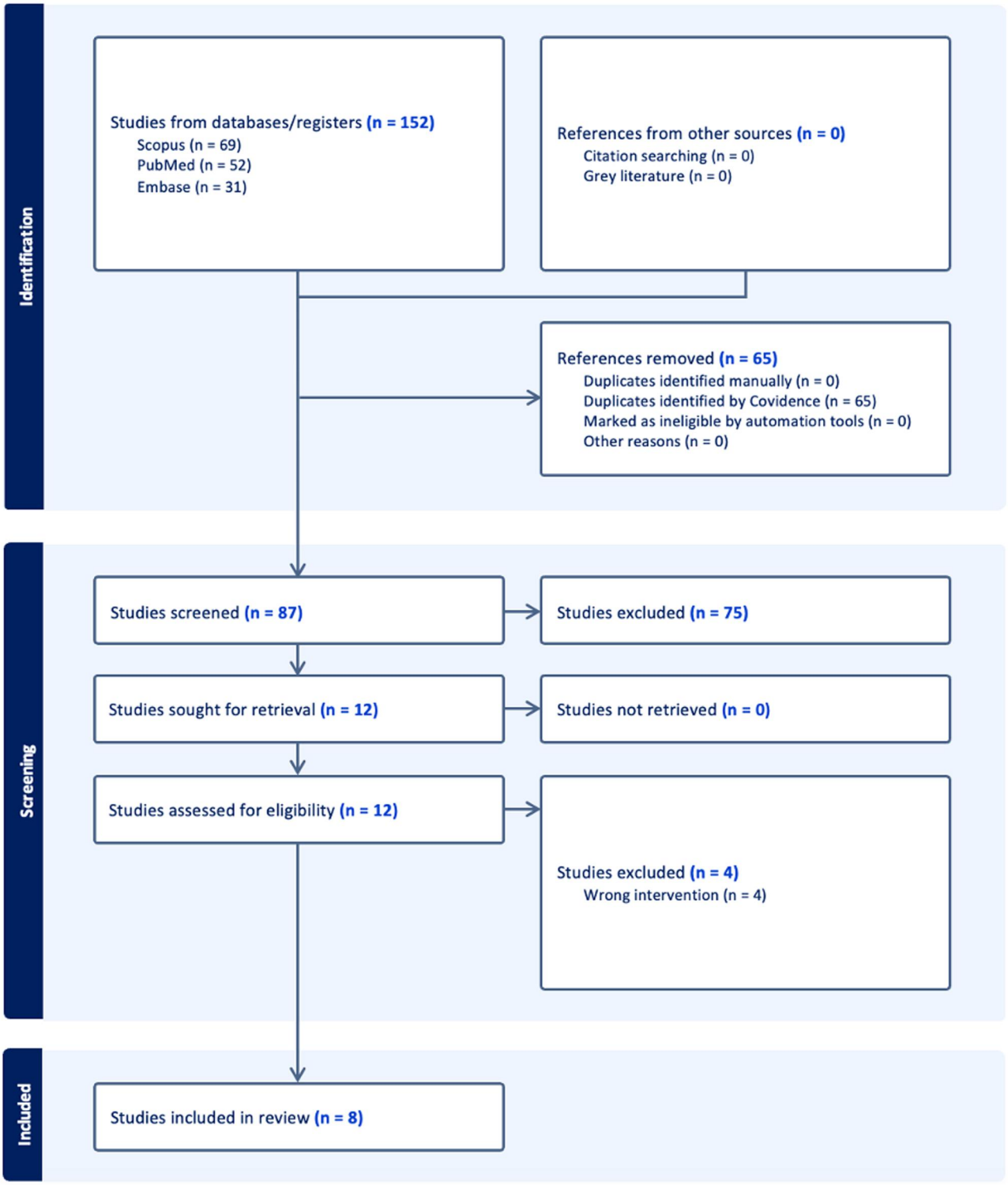

Figure 1. PRISMA flow diagram.

domain for Johnson et al, risk was found initially to be indeterminate. The study lacked clarity regarding enrolment method of the patients. Furthermore, there was a lack of exclusion criteria or information with respect to recruitment times and dates. Clarification was sought from the corresponding author regarding this, who clarified, and risk of bias was then found to be low.

Study characteristics

The summary table below (Table 2) offers an overview of eight studies. There was geographical diversity, with numbers of participants varying between 30 and 109. The age range of patients was 36-87 years, with a mean of 65-72 years, which is in-keeping with the epidemiology. No study had fewer than two radiologist readers. The range of experience was

Table 1. QUADAS-2 risk of bias results.

Publication	Risk of bias				Applicability		
	Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
Gassenmeier[23]	+	+	+	+	+	+	+
Gassenmeier[24]	+	+	+	+	+	+	+
Johnson et al[26]	+	+	+	+	+	+	+
Kim et al[27]	+	+	+	+	+	+	+
Lee et al[29]	+	+	+	+	+	+	+
Park et al[28]	+	+	+	+	+	+	+
Tong et al[30]	+	+	+	+	+	+	+
Ursprung et al[25]	+	+	+	+	+	+	+

1-14 years, where specified, with a mean of 5 years. One study[25] did not specify their radiologists' experience level.

All studies used 3T scanners, with Siemens MAGNETOM being the predominant subtype. All used T2WI and DWI in their acquisition. Studies looked at AI-acceleration in both T2WI and DWI,[29] only T2WI,[23,24,26,28,30] or only DWI.[25]

AI model characteristics

Various AI models were adopted for the eight studies, including convolutional neural networks (CNN), which have become the dominant AI models used in classification and detection-based radiological tasks, unrolled variational networks (UVV), variational neural networks (VNN), which overall resembles a UVV, research algorithms that were analogous to VNNs, and commercially available DL software. This variation, while indicative of a developing sector, presents challenges for this review. The heterogeneity complicates direct comparison of performance, efficacy, and clinical application when models differ in architecture, training, and underlying principles. Table 3 details the models used by each study.

Performance

Acquisition time

All studies found reductions in acquisition times to varying degrees. Gassenmaier et al[23] achieved a time reduction of 64.62% in axial imaging. A second study from the same author group[24] evaluated an alternate algorithm, identifying a reduction across all planes (64.62% for axial, 62.57% for coronal, and 60.51% for sagittal). Johnson et al demonstrated a reduction of 81.53% in axial, and 72.55% in coronal imaging. Kim et al saw reductions of 68.9% for high-resolution and 75.5% for low-resolution imaging. Lee et al saw reductions in the 32.5-32.6% range. Park et al saw reductions of 37.5% in tri-planar imaging. Tong et al reported a 69.91% reduction in axial imaging and a match between conventional and AI times in coronal imaging. The full details of acquisition times can be found in Supplementary Appendix A.

Qualitative measurements

All studies reported varying qualitative measurement, which can be found detailed in Supplementary Appendices B and C. Of the seven studies that looked at accelerated T2WI, all seven[23,24,26-30] reported overall imaging quality, four (57%)[23,24,27,29] reported artefact levels, three (43%)[23,24,27] reported noise levels, three (43%)[23,27,30] reported image sharpness, and two (29%)[23,24] reported diagnostic confidence.

Overall image quality was non-significant in three studies.[25,29,30] In one study,[28] it was significantly reduced in fast MRI, which was then significantly improved using DL reconstruction ($P < .001$). In the same study, using DL reconstruction of fast MRI maintained lesion delineation quality compared with conventional MRI ($P < .001$). Overall quality was significantly improved with AI acceleration ($P < .001$) in the first Gassenmaier et al[23] study. Discrete planes were reviewed for their second study,[24] with overall quality significantly improved or unchanged for axial (R1 and R2: $P < .001$), and improved for coronal (R1: $P = .002$ R2: $P < .001$) and sagittal (R1: $P = .002$ R2: $P = .005$). Kim et al found significant improvement in low-resolution AI compared to standard low-resolution MRI ($P = .001$).

Artefact levels were similar or improved compared to standard MRI in both Gassenmaier et al studies. The first study[23] showed statistical significance for both readers (R1: $P = .034$ and R2: $P = .011$). Their second study[24] showed similar significance across multiple planes (R1: $P = .003$ and R2: $P = .003$ for axial; R1: $P = .002$ and R2: $P = .014$ for coronal; R1: $P = .002$ and R2: $P = .011$ for sagittal). Conversely, both Kim et al and Lee et al found mostly statistical non-significance in artefact levels, apart from statistically significant lower scores for "*other (non-specified)*" artefact levels in AI-accelerated high-resolution ($P = .006$) in Kim et al.

Noise levels were improved with AI acceleration compared to standard MRI in both Gassenmaier et al studies. The first study[23] showed statistical significance for both readers (R1: $P < .001$ and R2: $P < .001$). Their second study[24] showed similar statistical significance across multiple planes (R1: $P < .001$ and R2: $P < .001$ for axial; R1: $P < .001$ and R2: $P < .001$ for coronal; R1: $P < .001$ and R2: $P < .001$ for sagittal). Kim et al also looked at noise levels in AI-accelerated low-resolution and high-resolution MRI and found that accelerated low-resolution MRI produced the same level of noise as conventional MRI, and this was statistically significant ($P < .001$). No other studies reported noise levels.

Image sharpness was improved with AI acceleration compared with standard MRI in both Gassenmaier et al studies. The first study[23] showed statistical significance for both readers (R1: $P < .001$ and R2: $P < .001$). Tong et al looked at sharpness of the capsule, PZ/TZ boundary, and periurethral area across both axial and coronal planes. They found statistically significant differences in clarity scores of the PZ/TZ boundary, and periurethral area for axial planes and the

Table 2. Study summary characteristics.

Study	Year	Country	Number of participants (following exclusions)	Mean/average age (range), in years	Number of readers	Reader experience (years)	Scanner type	MRI protocol	Position	Coils used
Gassenmaier et al[23]	2021	Germany	30	65 ± 11 (36-83)	2	3 years, 1 year	3T MAGNETOM Skyra and Prismafit	T2WI; DWI; ADC maps; T1WI; Contrast given	Supine	Combined 18 channel body; 12 elements of 32-channel spine
Gassenmaier et al[24]	2021	Germany	60	69 ± 9 (49-85)	2	8 years, 5 years	3T MAGNETOM Skyra, Prismafit, and Vida	T2WI; DWI; ADC maps; T1WI; Contrast given	Dorsal (supine)	Combined 18 channel body; 12 elements of 32-channel spine
Johnson et al[26]	2022	USA	30	68 ± 7	4	6 years, 3 years, 1 year, 1 year	3T MAGNETOM Vida	T2WI; DWI	Not specified	30 channel body coil
Kim et al[27]	2021	Korea	46	71 (71-87)	2	11 years, 7 years	3T MAGNETOM Vida	T2WI; DWI; T1WI; DCEI	Not specified	30 channel body coil; 32 or 72 channel spine coil
Lee et al[29]	2023	UK/Taiwan	40	68 (59-78)	2	14 years, 3 years	3T SIGNA Premier	T2WI; DWI	Not specified	30 channel anterior array coil; 60 channel spine coil
Park et al[28]	2022	Korea	109	67 (63-72)	3	> 10 years, 5 years, 2 years	3T GE Healthcare	T2WI; DWI; DCEI	Not specified	Not specified
Tong et al[30]	2023	USA	80	66 (47-84)	3	6 years, 4 years, 2 years	3T MAGNETOM Prisma	T2WI; DWI	Not specified	18 channel body overlaying pelvis
Ursprung et al[25]	2023	Germany	35	72 ± 7.5 (50-83)	3	Not specified	3T MAGNETOM Vida	T2WI; DWI; ADC maps; T1WI; Contrast given	Supine	18 channel body; 32 channel spine

Abbreviations: ADC = apparent diffusion coefficient; DCEI = dynamic contrast-enhanced imaging; DWI = diffusion weighted imaging; T1WI = T1-weighted imaging; T2WI = T2-weighted imaging.

Table 3. AI model and MRI scanner type.

Study	Year	MRI scanner	Image protocol	AI algorithm used
Gassenmaier et al[23]	2021	3T MAGNETOM Skyra and Prismafit	T2WI	Trainable convoluted neural network
Gassenmaier et al[24]	2021	3T MAGNETOM Skyra, Prismafit, and Vida	T2WI	Trainable unrolled variational network
Johnson et al[26]	2022	3T MAGNETOM Vida	T2WI	Variational neural network
Kim et al[27]	2021	3T MAGNETOM Vida	T2WI	Unrolled variational network
Lee et al[29]	2023	3T SIGNA Premier	T2WI	Commercial DL reconstruction software
Park et al[28]	2022	3T GE Healthcare	T2WI	Vendor convoluted neural network (AIR Recon DL, GE Healthcare)
Tong et al[30]	2023	3T MAGNETOM Prisma	T2WI; DWI	Commercial prototype (Siemens)
Ursprung et al[25]	2023	3T MAGNETOM Vida	T2WI; DWI; ADC maps; T1WI; Contrast given	DL reconstruction research application algorithm

Abbreviations: ADC = apparent diffusion coefficient; DWI = diffusion weighted imaging; T1WI = T1-weighted imaging; T2WI = T2-weighted imaging.

capsule, PZ/TZ boundary and periurethral area in coronal planes only with Reader 3 (axial PZ/TZ boundary $P = .003$; axial periurethral area $P = .03$; coronal capsule $P = .02$; coronal PZ/TZ boundary $P = .003$; coronal periurethral area $P = .01$). All other scores for both Reader 1 and Reader 2 were non-significant. Kim et al found no statistically significant differences in image sharpness. No other studies captured this image quality parameter.

Three studies (38%)[25,26,29] used AI acceleration on DWI and reported varying qualitative measurements. No study found any statistically significant difference in overall quality or artefacts between conventional DWI and AI-accelerated DWI. Similarly, neither Johnson et al nor Ursprung et al found any statistically significant difference in sharpness between conventional and AI-accelerated DWI.

Quantitative measurements

Four (50%)[25,27-29] studies report quantitative measurements, full details of which can be found in Supplementary Appendices D and E. Three studies[27-29] looked at AI-accelerated T2WI, whereas two[25,29] looked at DWI, either with T2WI or alone. Two T2WI studies[28,29] reported SNR and CNR. Park et al reported only SNR. For DWI, Lee et al reported both SNR and CNR, whereas Ursprung et al reported only SNR.

All three T2WI studies reported statistically significant differences in SNR between conventional and AI-accelerated MRI. Park et al reported significantly higher SNR for DL reconstructed fast MRI ($P < .001$). Conversely, Kim et al reported significantly decreased SNR in both the peripheral and transitional zones in AI-accelerated images ($P < .001$). Similarly, Lee et al reported statistically significant decreased SNR in high-resolution AI-accelerated imaging compared to conventional MRI ($P < .001$), however, they also reported significantly increased SNR when DL denoising was applied to images across all levels (low, medium, high) of denoising ($P < .001$).

Park et al reported statistically significant ($P < .001$) increased CNR for DL reconstructed fast MRI. Conversely, Lee et al found significantly decreased CNR in high-resolution AI-accelerated MRI protocol when compared with standard ($P < .001$). However, when DL denoising was applied to these images, there was statistically significant increased CNR ($P < .001$).

For DWI, Ursprung et al reported discretely reduced SNR in the TZ and obturator muscle in AI-accelerated DWI at $b = 1000$ s/mm^2 ($P < .001$), however SNR was similar across all other comparison sites and for all sites at $b = 0$ s/mm^2. Lee et al reported reduced SNR and CNR for AI-accelerated DWI at $b = 100$, 750, and 1400 s/mm^2 which were all statistically significant. However, again when denoising was applied both SNR and CNR were increased significantly at all levels and all b-values, with high levels of DL denoising showing the greatest effect size ($P < .001$).

Discussion

This systematic review highlights the developing research efforts in AI-accelerated prostate MRI, which have progressed since 2021. The research space is rapidly evolving and wide adoption of AI-acceleration is yet to take place, however, there is a sentiment of optimism across the literature. All studies showed notable reductions in acquisition times without compromising overall image quality. Similarly, the majority of the qualitative and quantitative measures captured by the studies showed consistency throughout AI-acceleration models when compared to conventional MRI protocols. However, there is variability in the results, which is perhaps attributable to the diverse methodologies, algorithms, and imaging protocols adopted across studies. The performance of AI is contingent on the robustness of the datasets with which it is trained. Efficacy, when trained on narrowly circumscribed and homogenous datasets, could be exemplary within a controlled setting, but may fail when challenged with diverse, real-world clinical or unfamiliar scenarios, which could account for the diverse results among these studies.

The promise of reducing acquisition time without compromising image quality stands as one of the most appealing facets. The apparent consistency in reductions in acquisition times holds potential to revolutionise practice by increasing patient throughput and improving comfort. It also presents an opportunity for healthcare systems to optimize resource allocation, potentially reducing wait times and

increasing access to MRI. The efficiency gains could translate into cost savings, making high-quality MRI more affordable and, by extension, more accessible to a broader segment of the population. Future research could explore AI's impact on healthcare economics and accessibility.

A highlight from both Gassenmaier et al studies is the observation that AI-acceleration either matched or improved upon artefact levels when compared with conventional MRI. The findings from both Kim et al and Lee et al indicate a predominantly neutral impact of AI on artefact levels. Studies that looked at artefact levels in DWI all found statistical non-significance, suggesting that the influence of AI on artefact levels may be algorithm, acquisition, or context specific.

Noise reduction could be instrumental in providing clarity in complex diagnostic scenarios. The equivalence in some studies suggests that noise reduction is potentially dependent on algorithm or imaging protocol. Several factors could account for these discrepancies, such as the training data, architecture, or optimization goals. Additionally, resolution could play a role, as high-resolution imaging tends to be more susceptible to noise due to smaller pixel size. While some found statistically significant improvements in image sharpness with AI, other studies presented a more tempered picture, once again underpinning the differences in algorithm. Others recorded significant variances in sharpness between readers, indicating that the effect may depend on interpretative experience.

Overall image quality incorporates many of the qualitative measurements discussed above. Some studies suggest a positive impact on overall quality; however, this is not consistent across the literature. Variability could suggest algorithms are optimized for certain resolutions and could under or overperform depending on context. A more nuanced reasoning is that image quality is inherently subjective and could introduce variability across studies or readers.

The varying SNR results across studies reflect the multifaceted impact of AI on image quality, along with underscoring the necessity to evaluate these changes within a clinical context. An enhancement in SNR potentially translates to clearer and more detailed images, thereby improving the diagnostic accuracy for clinicians, particularly in detecting subtle pathology. Lee et al observed that DL denoising can significantly increase SNR provides avenues for further optimization in AI protocols. Inherent variability in AI models, acquisition, and MRI technique all look to impact on SNR metrics. Improvements in CNR could lead to better differentiation between tissues, potentially significantly reducing diagnostic ambiguity and leading to more precise treatment planning. However, CNR saw similarly mixed results across the studies. Additionally, the application of denoising techniques by Lee et al may indicate that AI holds potential to adaptively enhance image contrast. However, solid conclusions cannot be made based on this evidence alone.

The varying SNR results across studies reflect the multifaceted impact of AI on image quality, along with underscoring the necessity to evaluate these changes within a clinical context. An enhancement in SNR potentially translates to clearer and more detailed images, thereby improving the diagnostic accuracy for clinicians, particularly in detecting subtle pathology. Improvements in CNR could lead to better differentiation between tissues, potentially significantly reducing diagnostic ambiguity and leading to more precise treatment planning.

Analysis also reveals notable differences between in SNR and CNR in T2WI and DWI. In T2WI, DL AI generally enhanced SNR, as evidenced by Park et al's findings, along with Kim et al and Lee et al who reported statistically significant decreased SNR without denoising, and statistically significant increased SNR with DL denoising, highlighting the potential of AI to either degrade or improve image quality based on the model's application and specifications. In DWI, the findings were even more mixed, with Ursprung et al observing reduced SNR in specific zones, whereas Lee et al documented a decrease in both SNR and CNR, which was subsequently improved with denoising. These results indicate that the impact on DWI is similarly dependent on the implementation of post-processing techniques, which can either exacerbate or mitigate the inherent challenges in DWI, such as its sensitivity to motion and diffusion noise.

The observed inconsistencies across studies likely stem from a combination of factors, including differences in architecture, training data quality and diversity, and specific parameters being processed. This highlights the importance of context and technique in evaluating AI's contribution to diagnostic imaging, underscoring the need for standardization in AI methodologies, which emerges as a key area needing rectification. The heterogeneity in metrics and methodologies across studies poses challenges in comparative appraisal and applicability. Establishing protocols and universal metrics would not only enhance comparative analysis but also ease translation into clinical practice. Similarly, training datasets require more scrutiny, as the efficacy is heavily dependent on quality and diversity. The emphasis must be channelled towards curating and leveraging datasets that are representative of a spectrum of clinical scenarios, ensuring the models are robust and versatile.

Limitations

One of the biggest limitations is heterogeneity in study designs. Variation in acquisition and AI models makes direct comparison complicated. Such disparity impedes the extrapolation and generalizability of results to broader contexts. Another limitation in these studies is the diverse profile of the radiologist readers that were used. Compounding this issue is the diverse profile of radiologist readers involved in these studies. Differences in training, experience, and the inherent subjectivity of some diagnostic metrics introduce a layer of reader bias that could significantly affect study outcomes. This variation underscores the necessity of considering reader experience in the interpretation of AI-accelerated images. Experienced radiologists may interpret these images differently from their less experienced counterparts, affecting diagnostic consistency and reliability. Lastly, an overarching limitation in the current body of research is temporal scope, both within the perspective of the studies themselves and the lack of long-term data overall.

This review itself has limitations. Due to the considerable heterogeneity, data were not pooled for formal meta-analysis however an extensive narrative synthesis that identified inconsistencies and limitations of the studies under review was included.

Conclusion

While the studies included in our review have provided some promising results, particularly around acquisition time, until larger-scale studies encompassing diverse clinical scenarios can confirm their findings, they should be considered as

preliminary. Further integration into clinical practice with the development of standardized evaluation metrics will be crucial to guide and inform future studies and clinical adaptation. Once widely adopted into clinical practice, AI-acceleration techniques hold potential to transform healthcare for practitioners, patients, and policymakers. AI-accelerated MRI modalities may also pave the way for a prostate screening programme. If validated in large-scale, longitudinal studies, it has the potential to fill the existing gap in early detection and transforming prognosis and management.

Author contributions

Ciaran Reinhardt (Methodology, Formal Analysis, Investigation, Data Curation, Writing—Original Draft Preparation, Writing—Review & Editing, Visualization), Hayley Briody (Data Curation, Writing—Review & Editing), and Peter J. MacMahon (Conceptualization, Writing—Review & Editing, Supervision)

Supplementary material

Supplementary material is available at *BJR* online.

Funding

None declared.

Conflict of interest

The authors declare no conflicts of interest.

References

1. Xiang D, Hu S, Mai T, et al. Worldwide cancer statistics of adults over 75 years old in 2019: a systematic analysis of the global burden of disease study 2019. *BMC Public Health*. 2022;22(1):1979. https://doi.org/10.1186/s12889-022-14412-1
2. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. *CA Cancer J Clin*. 2023;73(1):17-48. https://doi.org/10.3322/caac.21763
3. Gortz M, Huber AK, Linz T, et al. Detection rate of prostate cancer in repeat biopsy after an initial negative magnetic resonance imaging/ultrasound-guided biopsy. *Diagnostics (Basel)*. 2023;13(10):1761. https://doi.org/10.3390/diagnostics13101761
4. Klingebiel M, Arsov C, Ullrich T, et al. Data on the detection of clinically significant prostate cancer by magnetic resonance imaging (MRI)-guided targeted and systematic biopsy. *Data Brief*. 2022;45:108683. https://doi.org/10.1016/j.dib.2022.108683
5. Kuhl CK, Bruhn R, Kramer N, Nebelung S, Heidenreich A, Schrading S. Abbreviated biparametric prostate MR imaging in men with elevated prostate-specific antigen. *Radiology*. 2017;285(2):493-505. https://doi.org/10.1148/radiol.2017170129
6. 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(3):340-351. https://doi.org/10.1016/j.eururo.2019.02.033
7. Sarkar S, Das S. A review of imaging methods for prostate cancer detection. *Biomed Eng Comput Biol*. 2016;7(Suppl 1):1-15. https://doi.org/10.4137/BECB.S34255
8. Stabile A, Giganti F, Rosenkrantz AB, et al. Multiparametric MRI for prostate cancer diagnosis: current status and future directions. *Nat Rev Urol*. 2020;17(1):41-61. https://doi.org/10.1038/s41585-019-0212-4
9. Lin Y, Yilmaz EC, Belue MJ, Turkbey B. Prostate MRI and image quality: it is time to take stock. *Eur J Radiol*. 2023;161:110757. https://doi.org/10.1016/j.ejrad.2023.110757
10. Munn Z, Jordan Z. Interventions to reduce anxiety, distress and the need for sedation in adult patients undergoing magnetic resonance imaging: a systematic review. *Int J Evid Based Healthc*. 2013;11(4):265-274. https://doi.org/10.1111/1744-1609.12045
11. Oztek MA, Brunnquell CL, Hoff MN, et al. Practical considerations for radiologists in implementing a patient-friendly MRI experience. *Top Magn Reson Imaging*. 2020;29(4):181-186. https://doi.org/10.1097/RMR.0000000000000247
12. Hamilton J, Franson D, Seiberlich N. Recent advances in parallel imaging for MRI. *Prog Nucl Magn Reson Spectrosc*. 2017;101:71-95. https://doi.org/10.1016/j.pnmrs.2017.04.002
13. Yang D, Li C, Liu S, Liu Y. Compressed sensing image reconstruction based on convolutional neural network. *IJCIS*. 2019;12(2):873. https://doi.org/10.2991/ijcis.d.190808.001
14. Yang AC, Kretzler M, Sudarski S, Gulani V, Seiberlich N. Sparse reconstruction techniques in magnetic resonance imaging: methods, applications, and challenges to clinical adoption. *Invest Radiol*. 2016;51(6):349-364. https://doi.org/10.1097/RLI.0000000000000274
15. Johnson PM, Recht MP, Knoll F. Improving the speed of MRI with artificial intelligence. *Semin Musculoskelet Radiol*. 2020;24(1):12-20. https://doi.org/10.1055/s-0039-3400265
16. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts H. Artificial intelligence in radiology. *Nat Rev Cancer*. 2018;18(8):500-510. https://doi.org/10.1038/s41568-018-0016-5
17. Chu TN, Wong EY, Ma R, Yang CH, Dalieh IS, Hung AJ. Exploring the use of artificial intelligence in the management of prostate cancer. *Curr Urol Rep*. 2023;24(5):231-240. https://doi.org/10.1007/s11934-023-01149-6
18. Sushentsev N, Moreira Da Silva N, Yeung M, et al. Comparative performance of fully-automated and semi-automated artificial intelligence methods for the detection of clinically significant prostate cancer on MRI: a systematic review. *Insights Imaging*. 2022;13(1):59. https://doi.org/10.1186/s13244-022-01199-3
19. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71. https://doi.org/10.1136/bmj.n71
20. Mongan J, Moy L, Kahn CE Jr. Checklist for artificial intelligence in medical imaging (CLAIM): a guide for authors and reviewers. *Radiol Artif Intell*. 2020;2(2):e200029. https://doi.org/10.1148/ryai.2020200029
21. Innovation VH. Covidence Systematic Review Software. Accessed 7th August, 2023.
22. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529-536. https://doi.org/10.7326/0003-4819-155-8-201110180-00009
23. Gassenmaier S, Afat S, Nickel D, Mostapha M, Herrmann J, Othman AE. Deep learning-accelerated T2-weighted imaging of the prostate: Reduction of acquisition time and improvement of image quality. *Eur J Radiol*. 2021;137:109600. https://doi.org/10.1016/j.ejrad.2021.109600
24. Gassenmaier S, Afat S, Nickel MD, et al. Accelerated T2-weighted TSE imaging of the prostate using deep learning image reconstruction: a prospective comparison with standard T2-weighted TSE imaging. *Cancers (Basel)*. 2021;13(14):3593. https://doi.org/10.3390/cancers13143593
25. Ursprung S, Herrmann J, Joos N, et al. Accelerated diffusion-weighted imaging of the prostate using deep learning image reconstruction: a retrospective comparison with standard diffusion-weighted imaging. *Eur J Radiol*. 2023;165:110953. https://doi.org/10.1016/j.ejrad.2023.110953

26. Johnson PM, Tong A, Donthireddy A, et al. Deep learning reconstruction enables highly accelerated biparametric MR imaging of the prostate. *J Magn Reson Imaging*. 2022;56(1):184-195. https://doi.org/10.1002/jmri.28024
27. Kim EH, Choi MH, Lee YJ, Han D, Mostapha M, Nickel D. Deep learning-accelerated T2-weighted imaging of the prostate: impact of further acceleration with lower spatial resolution on image quality. *Eur J Radiol*. 2021;145:110012. https://doi.org/10.1016/j.ejrad.2021.110012
28. Park JC, Park KJ, Park MY, Kim MH, Kim JK. Fast T2-weighted imaging with deep learning-based reconstruction: evaluation of image quality and diagnostic performance in patients undergoing radical prostatectomy. *J Magn Reson Imaging*. 2022;55(6):1735-1744. https://doi.org/10.1002/jmri.27992
29. Lee KL, Kessler DA, Dezonie S, et al. Assessment of deep learning-based reconstruction on T2-weighted and diffusion-weighted prostate MRI image quality. *Eur J Radiol*. 2023;166:111017. https://doi.org/10.1016/j.ejrad.2023.111017
30. Tong A, Bagga B, Petrocelli R, et al. Comparison of a deep learning-accelerated vs. conventional T2-weighted sequence in biparametric MRI of the prostate. *J Magn Reson Imaging*. 2023;58(4):1055-1064. https://doi.org/10.1002/jmri.28602

British Journal of Radiology, 2024, 97, 1234–1242
https://doi.org/10.1093/bjr/tqae093
Systematic Review