Radiography 32 (2026) 103380

Contents lists available at ScienceDirect

Radiography

journal homepage: www.elsevier.com/locate/radi

Visual assessment of AI-reconstructed knee MRI: A pilot study

S.S. Ghotra [a,b,c,*], L. Buttex [b], L. Gallus [d], Y. Cottier [e], J. McNulty [c], A. McGee [c], C. Sá dos Reis [a]

[a] School of Health Sciences (HESAV), University of Applied Sciences and Arts Western Switzerland (HES-SO), Lausanne 1011, Switzerland
[b] Department of Diagnostic & Interventional Radiology, Lausanne University Hospital (CHUV) and University of Lausanne (UNIL), Lausanne 1011, Switzerland
[c] Radiography and Diagnostic Imaging, School of Medicine, University College Dublin, Ireland
[d] Department of Radiology, Jura Hospital, Delémont, Switzerland
[e] Centre d'Imagerie Diagnostique de Lausanne, Lausanne 1011, Switzerland

ARTICLE INFO

Article history:
Received 8 January 2026
Received in revised form
13 February 2026
Accepted 16 February 2026
Available online xxx

Keywords:
Deep learning
Image reconstruction
Magnetic resonance imaging
Musculoskeletal
Artificial intelligence

STRUCTURED ABSTRACT

Introduction: Artificial intelligence is increasingly influencing medical imaging workflows by enhancing image quality and reducing acquisition time. The purpose of this study is to evaluate the use of artificial intelligence (AI) reconstruction methods for knee magnetic resonance imaging (MRI) investigations.
Methods: An exploratory comparison was performed between AI-enhanced and standard 3T knee MRI protocols. Sequences included sagittal T1-weighted (T1w) fast spin echo (FSE), proton density-weighted (PDw) FSE fat saturation (FS) and T2w FSE, with additional PDw FSE FS axial and coronal views. Parameters were adjusted to improve image quality (IQ) and shorten scan duration. The optimised AI protocol was tested on ten healthy volunteers. Three MRI experts independently assessed images visually using ViewDEX. Visual grading analysis (VGA), inter-observer agreement (Kappa), and visual grading characteristics (VGC) were utilised for evaluation.
Results: VGA results demonstrated that AI reconstruction produced equal or superior scores across most anatomical and IQ criteria, with a time reduction of 36.9 % (8:22 min vs 13:15 min). All the AI-enhanced sequences were judged as clinically acceptable. Kappa values indicated moderate-to-good agreement for AI sequences, whereas agreement for standard sequences ranged from low to good. VGC confirmed statistically higher performance for AI images (AUC_{VGC} 0.76–0.81, $p \leq 0.05$).
Conclusion: This study indicates that incorporating AI into knee MRI protocols can substantially enhance overall image quality while simultaneously reducing acquisition time by 36.9 %. However, further research is needed to reinforce these findings through clinical validation.
Implication for practice: AI can reduce scan time without lowering image quality, improving workflow and patient throughput. Understanding how acquisition parameters affect image quality and artefacts is crucial when integrating AI reconstruction into protocols.

Introduction

The knee is a complex joint composed of bones, ligaments, tendons, bursae, and other soft tissues, all of which may be affected by pathologies such as tendinitis, bursitis, ligament injuries, trauma, or osteoarthritis.[1] Knee pain represents a frequent clinical issue, accounting for roughly 6 % of primary care consultations in Switzerland, with nearly half of the population reporting this at least once in their lifetime.[2]

Multiple imaging modalities can be used to assess the knee, including projection radiography, computed tomography (CT), magnetic resonance imaging (MRI), and ultrasound. Radiography and CT expose patients to ionising radiation and have limited soft-tissue contrast, while ultrasound struggles with deep and intra-osseous structures due to bone interference.[3] MRI has therefore become the gold standard for evaluating soft tissue structures of the knee joint.[4,5] However, it remains costly, less accessible, and time-consuming, potentially reducing patient comfort,[6,7] especially in anxious or claustrophobic individuals,[8,9] and presents

* Corresponding author. School of Health Sciences (HESAV), University of Applied Sciences and Arts Western Switzerland (HES-SO), Lausanne 1011, Switzerland.
E-mail addresses: switindersingh.ghotra@hesav.ch (S.S. Ghotra), lucas.buttex@chuv.ch (L. Buttex), leo.gallus@h-ju.ch (L. Gallus), yann.cottier@hotmail.ch (Y. Cottier), jonathan.mcnulty@ucd.ie (J. McNulty), allison.mcgee@ucd.ie (A. McGee), claudia.sadosreis@hesav.ch (C. Sá dos Reis).

https://doi.org/10.1016/j.radi.2026.103380

safety contraindications related to magnetic fields and implanted devices.[10,11]

A typical knee MRI scan lasts 15–20 min,[12] increasing vulnerability to motion artefacts in patients who cannot remain still, potentially compromising diagnosis.[13] Image quality depends on spatial resolution, contrast, noise, and artefacts.[14,15] To address acquisition time and quality limitations, artificial intelligence (AI) has emerged as a promising solution,[16–19] enabling faster scans followed by deep-learning reconstruction to enhance signal-to-noise ratio (SNR) or spatial resolution.[16,17,20]

These methods leverage deep learning models trained on raw or reconstructed data to enhance spatial resolution, suppress noise, and preserve anatomical detail. Within this context, vendor-specific implementations, such as deep learning–based sharpening and denoising frameworks, have shown the ability to achieve high SNR and consistent anatomical reproduction while enabling reduced acquisition times.[21–27]

Despite its potential, AI performance varies depending on clinical use-cases,[28,29] requires suitable hardware,[30] and may amplify or create artefacts unrecognised in algorithm training.[21,31] Proper understanding by MRI staff is essential for safe clinical implementation.[32,33] Although AI promises workflow benefits, literature remains limited regarding comparative image-quality evaluation and potential hallucination risks in clinical settings.[21,34] The main objective of this study is to evaluate the use of AI reconstruction methods for knee MRI investigations.

Methods

Study design

This pilot exploratory study was conducted in three main steps: Image acquisition, Image evaluation, and Data analysis.

Study data were acquired prospectively after informing and obtaining written consent from each volunteer participant, in accordance with ethical principles. All data were anonymised, safely stored in a secured server with limited access to the authors and have been managed in compliance with the General Data Protection Regulation (GDPR). Ethical approval for this study has been granted by the Ethics Committee Vaud (CER-VD, Project-ID: 2024-01679).

Image acquisition

Study population

For this pilot study, 10 healthy volunteers, aged 18–35 years, were recruited regardless of their gender or body mass index (BMI), in order to take into account the diversity of the population and to assess the performance of the sequences in a diverse population.[35] The age criterion was chosen based on the completion of growth and the general good health of the knee. From a certain age, the joint and the bone structure of the knee are more susceptible to wear.[36,37] Participants with known knee pathologies, contraindications to MRI, claustrophobia or pregnancy were excluded. Participation was entirely voluntary. The research information sheet and the consent form were provided to the participants in advance. They were given at least one week to reflect and decide whether to participate in this study.

MRI scan protocol

Image acquisitions were performed on a 3T MAGNETOM Vida Fit with XA60 software (Siemens Healthineers, Erlangen, Germany) in December 2024 using a dedicated 15-channel transmit–receive knee coil. Multiplanar fast spin echo (FSE) T1-weighted (T1w), T2-weighted (T2w), and proton density-weighted (PDw) sequences with chemical shift-selective (CHESS) fat saturation (FS) were acquired with and without the manufacturer's AI solution (Table 1). Siemens Deep Resolve (DR) accelerates 2D FSE MRI, with Deep Resolve Boost (DRB) and Deep Resolve Gain (DRG) applying raw-data–based iterative denoising to increase SNR. DRB is currently restricted to 2D FSE with short tau inversion recovery (STIR) or CHESS, whereas DRG can also be combined with DIXON (a chemical shift imaging method facilitating water and fat signal separation). Denoising strength is adjustable (DRB: low/medium/high; DRG: denoising levels 1/4/8 and edge-enhancing factor of 2), and both operate with Deep Resolve Sharp (DRS) for sharpness enhancement.[38,39] By reducing noise and improving spatial resolution, Deep Resolve enables the acquisition of lower spatial resolution and higher-noise images with modified parameters, shortening scan time. However, such adjustments may also alter contrast or increase smoothing and artefacts, affecting anatomical fidelity.[31]

Knee parameters modification

The first optimisation phase was performed on one participant. Acquisition parameters were adjusted by modifying spatial resolution, parallel imaging acquisition, reference scan type (integrated vs. separate), and oversampling, in addition to testing different AI reconstruction levels. The best DRB protocol meeting clinical image quality requirements, as determined by the researcher/independent assessor(s) against defined image quality criteria, was then validated against the standard protocol in 10 participants.

During optimisation, multiple configurations were compared using DRB and DRB + DRS. Parameters tested included: 1) spatial resolution reductions of 10 %, 30 %, and 50 %; 2) integrated Parallel Acquisition Techniques (iPAT factor 3 or 4); 3) integrated vs. non-integrated FSE reference scans; and 4) oversampling levels of 50 %, 100 %, and 200 %.[40] Each adjustment affected acquisition time, spatial resolution, and SNR. Final parameters used for the standard (STD) and Deep Resolve Boost (DRB) sequences acquired for this study are shown in Table 2.

Image evaluation

Visual grading analysis (VGA) was the approach used to assess images with human observers as the evaluation concerned the visibility and reproduction of healthy anatomical structures, as well as image quality.[41] VGA provides detailed information on the performance of a new technique, as well as for optimising image

Table 1
Knee MRI initial test acquisition protocol.

Standard protocol	Deep Resolve Gain and Sharp protocol	Deep Resolve Boost and Sharp protocol
T1w FSE sagittal	T1w FSE sagittal	T1w FSE sagittal
PDw FSE FS sagittal	PDw FSE FS sagittal	PDw FSE FS sagittal
T2w FSE sagittal	T2w FSE sagittal	T2w FSE sagittal
PDw FSE FS axial	PDw FSE FS axial	PDw FSE FS axial
PDw FSE FS coronal	PDw FSE FS coronal	PDw FSE FS coronal

Table 2
MRI acquisition parameters.

Imaging parameters	PDw FS SAG FSE		PDw FS COR FSE		PDw FS AX FSE		T1w SAG FSE		T2w SAG FSE	
	STD	DRB	STD	DRB	STD	DRB	STD	DRB	STD	DRB
TR repetition time (msec)	3720	3720	4940	4210	4360	4610	786	754	5670	3100
TE echo time (msec)	35	35	25	37	38	36	9.4	9	87	76
Flip angle	150	150	180	180	150	150	123	120	150	159
Slice thickness (mm)	3	3	3	3	3	3	3	3	3	3
Acquisition voxel size (mm)	0.59 × 0.44 × 3	0.59 × 0.44 × 3	0.47 × 0.38 × 3	0.60 × 0.48 × 3	0.54 × 0.38 × 3	0.58 × 0.41 × 3	0.47 × 0.38 × 3	0.54 × 0.41 × 3	0.71 × 0.53 × 3	0.51 × 0.41 × 3
Reconstruction voxel size (mm)	0.22 × 0.22 × 3	0.22 × 0.22 × 3	0.38 × 0.38 × 3	0.24 × 0.24 × 3	0.38 × 0.38 × 3	0.20 × 0.20 × 3	0.19 × 0.19 × 3	0.20 × 0.20 × 3	0.53 × 0.53 × 3	0.20 × 0.20 × 3
Field of view (mm^2)	170 × 170	170 × 170	170 × 170	170 × 170	170 × 159	170 × 159	170 × 170	170 × 170	170 × 170	171.6 × 170
FOV phase (%)	100 %	100 %	100 %	100 %	93.8 %	93.8 %	100 %	100 %	100 %	101 %
Acquisition matrix	384 × 288p	384 × 288p	448 × 358p	352 × 282p	448 × 294p	416 × 273p	448 × 358p	416 × 312p	320 × 240p	416 × 336p
Fourier plan sampling	100 %	100 %	100 %	100 %	100 %	100 %	100 %	100 %	100 %	100 %
Encoding phases	Head-feet	Head-feet	Head-feet	Head-feet	Right-left	Right-left	Head-feet	Head-feet	Head-feet	Head-feet
AI reconstruction method		DRB medium DRS		DRB medium DRS		DRB medium DRS		DRB medium DRS		DRB medium DRS
Parallel imaging type	GRAPPA	GRAPPA	GRAPPA	GRAPPA	GRAPPA	GRAPPA	GRAPPA	GRAPPA	GRAPPA	GRAPPA
Parallel imaging factor	2	4	2	4	2	4	2	4	2	4
Turbo factor	7	7	7	8	7	8	3	3	14	14
Number of excitations	1	1	1	1	1	1	1	1	1	1
Bandwidth (Hz/pixel)	303	303	199	200	328	325	243	240	220	236
Acquisition time	03:02	02:00	03:32	01:31	02:02	01:21	03:03	01:24	01:36	02:06

quality and protocols. Observers rate their personal perception and their confidence in visualising the defined anatomical structures, as well as the image quality.

For this type of study, absolute VGA was chosen, as the objective was to evaluate the images singularly rather than compare them directly to a reference image, which would require the use of relative VGA.[41]

Criteria for evaluating knee MRI images

Six main anatomical criteria were compiled from the literature.[42–45] The list included various anatomical structures necessary for the evaluation of the knee joint. Several criteria were included to assess image quality and the presence of artefacts or distortions. Finally, one question evaluated whether observers were confident in the clinical acceptability of the images they reviewed (Table 3).

Observers assessed the images according to the anatomical and image quality criteria using a grading scale based on a four-point Likert scale. Anatomical criteria were graded as 1 ″not defined", 2 "partially defined", 3 "defined", 4 "clearly defined", while image quality criteria were graded as 1 "very poor", 2 "poor", 3 "adequate", 4 "very good". They also had to indicate whether an artefact or image distortion affected anatomical structure visualisation with a closed-end question, yes/no.

Conditions for image evaluation

Image evaluation was performed under clinical viewing conditions (≤50 lux) using the same workstation and diagnostic monitor for all readers. Images were assessed in ViewDEX 3.0 ((Viewer for Digital Evaluation of X-ray Images), a Java-based DICOM viewer for observer studies.[46–49] The software enabled scrolling through full image stacks without allowing revision of previous scores, and individual login accounts ensured separate data collection for each observer.

Subjective evaluation by observers

Images were evaluated by musculoskeletal (MSK) MRI experienced radiographers with 16–20 years of clinical experience, to ensure anatomical and clinical image-quality adequacy. The study protocol and software use were explained beforehand to standardise image assessment. A known limitation of VGA is the absence of a reference image, which can increase intra- and inter-observer variability,[41] this was mitigated by clearly outlining the evaluation objectives and criteria prior to evaluate.

Data analysis

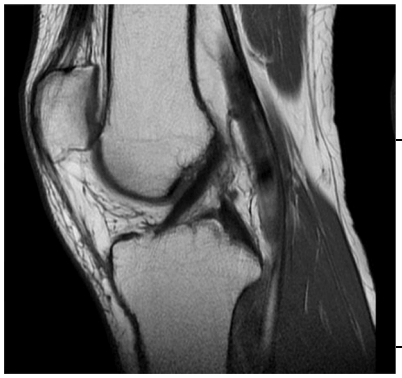
Visual grading analysis

Mean scores were calculated for each sequence based on descriptive anatomical criteria to determine whether anatomical definition was clinically adequate (mean > 3) and images were clinically acceptable. Each criterion was rated on a four-point Likert scale, and the presence artefacts/distortions affecting critical anatomy was recorded via a yes/no response. This allowed evaluation of whether a sequence met clinical image-quality requirements, but not direct comparison of performance, as Likert scales lack uniform interval distances and are subject to perceptual variability.[41]

Visual Grading Characteristics

As absolute VGA is limited to mean ordinal scores and cannot reliably identify superior sequences, image quality comparisons were performed using the validated Visual Grading Characteristics (VGC) method.[41,50,51] VGC applies non-parametric analysis to ordinal data without assuming a specific distribution and

Table 3
Image evaluation criteria.

Knee MRI		Anatomical Criteria				
	1	Bone structures are:	1	2	3	4
	2	Bone marrow is:	1	2	3	4
	3	Cortical bone is:	1	2	3	4
	4	The meniscus is:	1	2	3	4
	5	The ligaments are:	1	2	3	4
	6	The tendons are:	1	2	3	4
	Image quality criteria					
	7	Overall image quality is:	1	2	3	4
	8	The noise is	1	2	3	4
	9	The contrast is:	1	2	3	4
	10	The fluid-cartilages contrast is	1	2	3	4
	11	The fluid-soft tissue contrast is	1	2	3	4
	12	The definition (resolutions) of the structure is	1	2	3	4
	13	The blur is:	1	2	3	4
	Artefact and distortion					
	14	Does any artefact prevent correct visualisation of the anatomy?	Yes		No	
	15	Does image distortion prevent correct visualisation of the anatomy?	Yes		No	
	Clinical acceptability					
	16	Would you clinically accept these images?	1	2	3	4

Anatomical criteria (1–6) are assessed on a Likert scale of 1–4: 1 'not defined'; 2 'partially defined'; 3 'defined'; 4 'clearly defined'.
Image quality criteria (7–13) are assessed on a Likert scale of 1–4: 1 'very poor'; 2 'poor'; 3 'adequate'; 4 'very good'.
Problems with artefacts, distortion for relevant anatomy (14–15) are assessed by a closed 'yes' or 'no' response question.
Clinical acceptability (16) is assessed on a Likert scale of 1–4: 1 'not confident'; 2 'partially confident'; 3 'confident'; 4 'clearly confident'.

accounts for systematic observer ranking. As VGC is limited to pairwise comparisons, multiple variables were analysed using separate VGC analyses with VGC Analyser software.[52–54]

Differences in image quality were quantified using the area under the VGC curve (AUC_{VGC}). An AUC_{VGC} of 0.5 indicates equivalent image quality, values > 0.5 favour the test sequence, and values < 0.5 favour the reference sequence.[50,51]

Inter-observer agreement

Intra-observer agreement was assessed using the Fleiss' kappa statistic, calculated from categorised VGA scores and interpreted according to Landis and Koch.[55–57]

Results

Participants' characteristics

10 participants were enrolled, seven women and three men. Age ranged mainly from 22 to 25 years, with an outlier participant of 30 years, weight from 55 to 88 kg, mainly with an outlier of 125 kg, height from 162 to 187 cm, and body mass index (BMI) from 20.07 to 27.68 (kg/cm^2), with an outlier of 40.35 BMI(Table 4).

Table 4
Participant's characteristics.

Number	Gender	Age	Weight (kg)	Height (cm)	BMI (kg/cm2)
1	M	25	125	176	40.35
2	F	30	66	171	22.57
3	F	22	55	162	20.96
4	F	24	68	162	25.91
5	M	22	85	187	24.31
6	M	24	83	178	26.20
7	F	22	80	170	27.68
8	F	23	60	170	20.76
9	F	25	68	173	22.72
10	F	24	58	170	20.07

Test and optimisation phase

Multiple parameter combinations were tested to optimise image quality while reducing acquisition time, specific absorption rate (SAR), and artefacts. The final protocol selected (Table 2) incorporated DRB medium, iPAT 4, a separate reference scan, 10 % spatial resolution reduction, which was achieved by reducing the matrix size and therefore increasing pixel size at a constant field-of-view (FOV), and 100 % oversampling. The total time of the final protocol selected (08:22) is fast than the standard protocol (13:15), resulting in 36.9 % reduction in acquisition time.

DRG produced noisier images than DRB (Fig. 1). Strong DRB introduced smoothing, whereas weak DRB increased noise, making medium DRB the most balanced. iPAT 3 offered no SNR benefit and increased scan time compared with iPAT 4. Integrated reference scans reduced anatomical detail and increased noise. Simultaneous Multi-Slice (SMS) increased SAR. Matrix reductions of 30–50 % degraded spatial resolution (bigger pixel size) and introduced blurring relative to a 10 % reduction (Fig. 2). Oversampling at 50 % caused artefacts, while 150 % prolonged scan time without quality improvement over 100 % oversampling (Fig. 3).

Visual grading analysis

For the sagittal PDw FS FSE sequence, VGA scores were consistently higher with DRB than with the corresponding standard. DRB scored highest for the meniscus (3.77) and lowest for bone (3.57), while standard achieved its best VGA scores for fluid–cartilage and fluid–soft tissue contrast (3.60) and its lowest for sharpness (2.60). Artefacts, almost not present, were mainly reported in DRB than in the standard (Fig. 5).

For the coronal PDw FS FSE sequence, DRB sequence also outperformed the standard protocol, with the highest score for fluid–soft tissue contrast (3.93) and the lowest for overall image quality (3.43). In comparison, the standard sequence reached 3.30 for fluid–soft tissue and fluid–cartilage contrast, with its lowest score in sharpness (2.30) (Fig. 6).

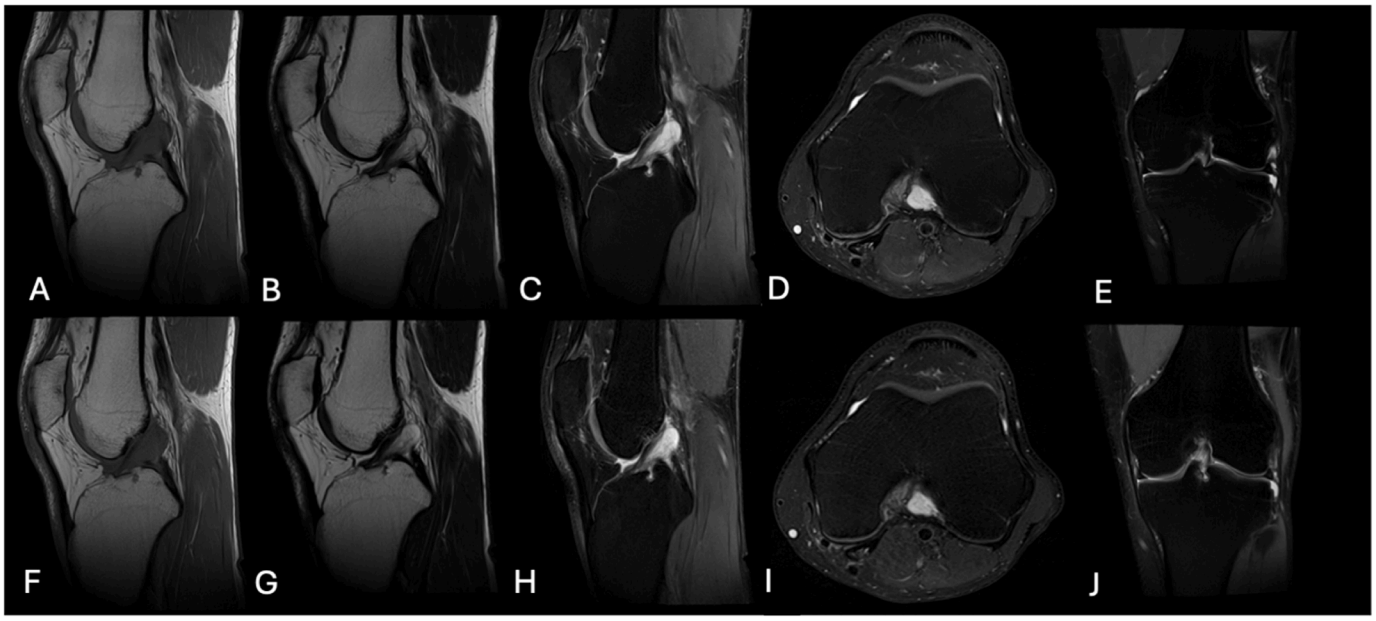


Figure 1. Comparison of the protocol with deep resolve boost and sharp on the first line (A. T1w FSE sagittal, B. T2w FSE sagittal, C. PDw FS sagittal, D. PDw FS axial and E. PDw FS coronal) and the protocol with deep resolve gains and sharp on the second line (F. T1w FSE sagittal, G. T2w FSE sagittal, H. PDw FS sagittal, I. PDw FS axial and J. PDw FS coronal).

For the axial PDw FS FSE sequence, DRB was again rated higher overall than STD, with the highest score for fluid–soft tissue contrast (3.77) and the lowest for overall image contrast (3.37). In the standard protocol, overall contrast was slightly higher than DRB (3.40 vs 3.30), while the best score reached 3.63 for fluid–soft tissue contrast and the lowest was 2.50 for image noise (Fig. 7).

For the sagittal T1w FSE sequence, DRB achieved the study's highest score (4.00) for cortical bone and a lowest score of 3.63 for fluid–soft tissue contrast. In comparison, standard scored highest for noise (3.70) and lowest for sharpness (2.90). Overall, DRB outperformed the standard across all evaluated criteria (Fig. 8).

For the sagittal T2w FSE sequence, the highest score was 3.97 for tendons, with the lowest at 3.63 for both fluid–cartilage and fluid–soft tissue contrast. In the standard protocol, noise scored highest (3.50) and sharpness lowest (2.03), notably below the DRB sharpness score (3.87). This sequence demonstrated the greatest performance gap between DRB and standard across all sequences (Fig. 9).

Overall, VGA scores show that DRB sequences were equal or superior to standard sequences for most criteria, except sagittal T2w, PDw FS, and axial PDw FS, where artefacts were more frequent in DRB sequences. Despite this, DRB was generally rated higher in perceived image quality and provided a 36.9 % reduction in acquisition time (Table 2 and Fig. 4).

Visual grading characteristics (VGC)

Regarding the VGC analysis, the vertical axis of the graph represents the test condition, meaning the DRB sequences, and the horizontal axis represents the reference conditions, meaning the standard sequences (Fig. 10). The DRB is performing significantly better compared to the standard protocol (p-value<0.05; 95%CI) with AUC_{VGC} values ranging from 0.76 to 0.81, indicating good discrimination in favour of DRB sequences (Table 5).

Inter-observer agreement

For DRB sequences, the results of the kappa test ranged from 0.35 to 0.71, whereas for standard sequences, they ranged from 0.12 to 0.71. On average, DRB sequences achieved a score of 0.52,

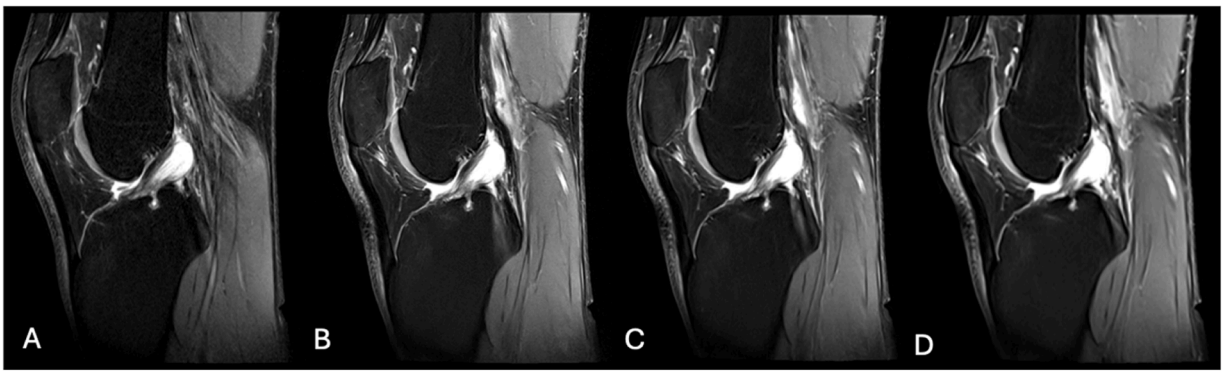


Figure 2. Comparison of the PDw FS sagittal sequences with matrix reduction while using deep resolve boost and deep resolve sharp image reconstruction. 1. PDw FSE sagittal no AI reconstruction, 2. PDw FSE sagittal with DRB & DRS and 10 % matrix reduction, 3. PDw FSE sagittal with DRB & DRS and 30 % matrix reduction, 4. PDw FSE sagittal with DRB & DRS and 50 % matrix reduction.

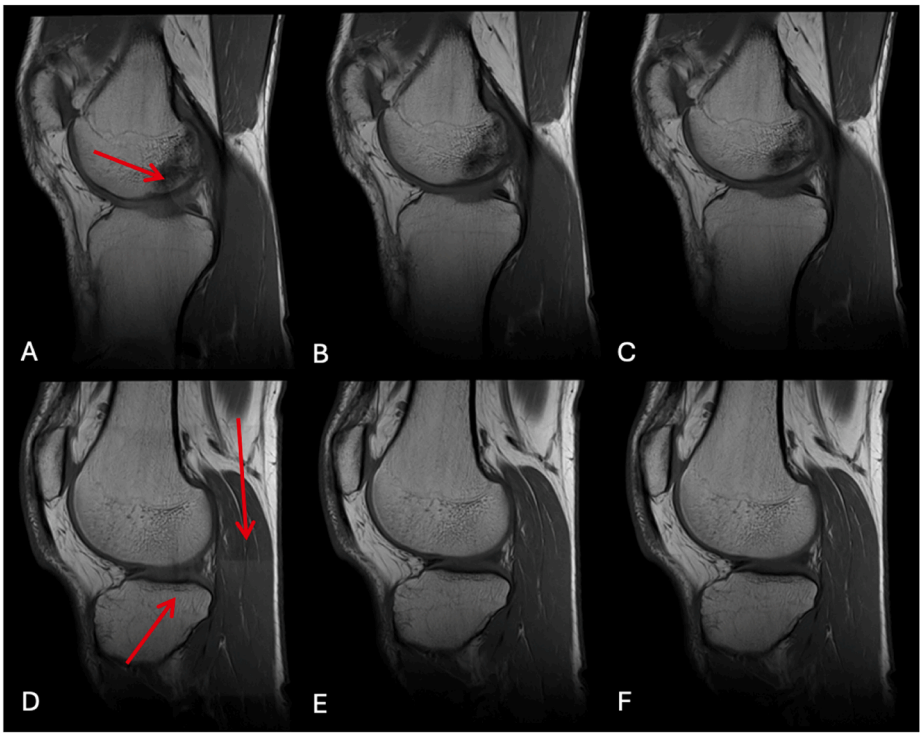


Figure 3. Comparison of the T1 sagittal sequences with variable percentage oversampling while using deep resolve boost and sharp image reconstruction. 1. and 4. PDw FSE sagittal AI 50 % oversampling, 2. and 5. PDw FSE sagittal AI 100 % oversampling, 3. and 6. PDw FSE sagittal AI 150 % oversampling. Artefacts are seen only on at 50 % oversampling acquisition (red arrows).

corresponding to a "moderate" level of agreement. In contrast, the average for standard sequences was 0.31, corresponding to a "fair" level of agreement. These results indicate that observers tended to agree more often when evaluating DRB sequences (Table 6).

Discussion

The main objective of this study was to evaluate the use of AI in knee MRI examinations, and to this end, two sub-objectives were defined.

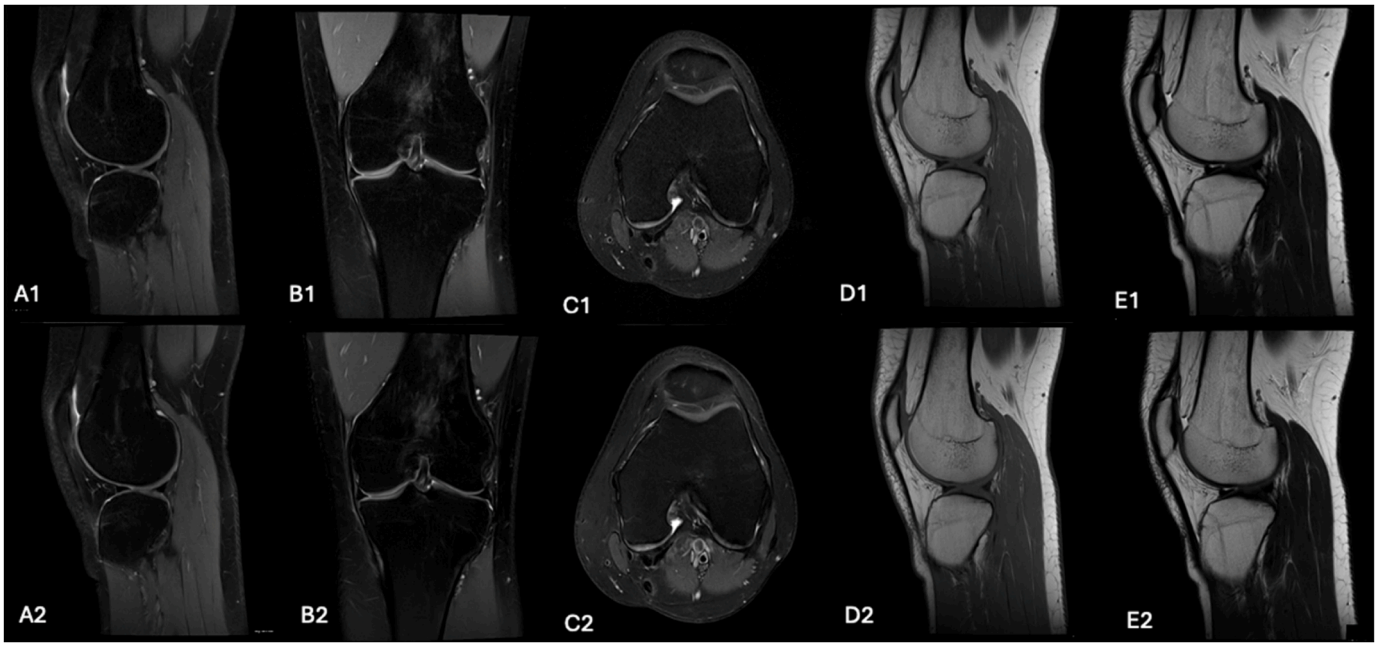


Figure 4. PDw FS sagittal FSE: A1) standard (STD) and A2) deep resolve boost (DRB) and sharp (DRS); PDw FS coronal FSE: B1) STD and B2) DRB and DRS; PDw FS axial FSE: C1) STD and C2) DRB and DRS; T1w sagittal FSE: D1) STD and D2) DRB and DRS; T2w sagittal FSE: E1) STD and E2) DRB and DRS.

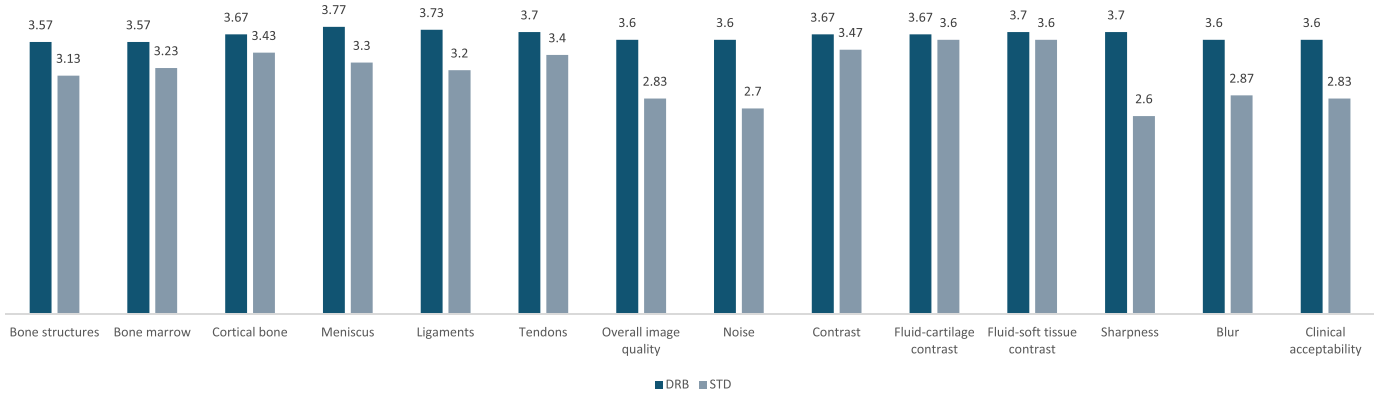


Figure 5. Visual grading analysis of Sagittal PDw FS FSE.

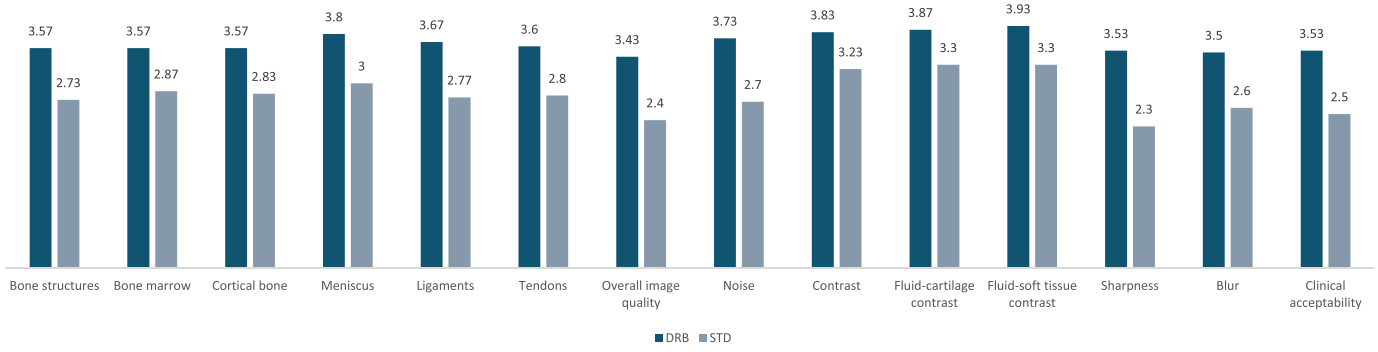


Figure 6. Visual grading analysis of Coronal PDw FS FSE.

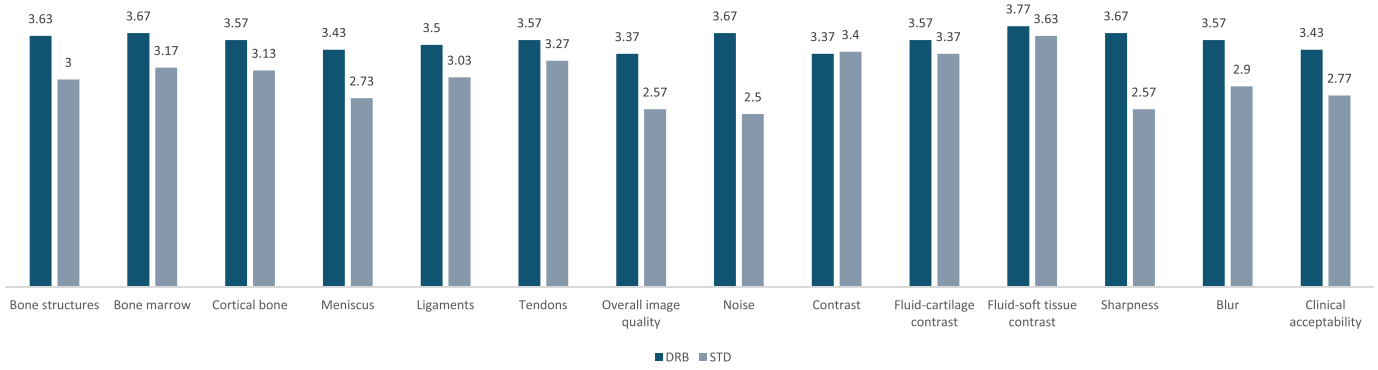


Figure 7. Visual grading analysis of Axial PDW FS FSE.

Overall, DRB sequences consistently matched or outperformed standard sequences in VGA scoring, indicating superior image quality alongside a 36.9 % reduction in acquisition time. DRB produced higher and more uniform scores across structures (bone, bone marrow, meniscus, ligaments, tendons), with no criterion falling below the clinical acceptability threshold of 3, unlike standard, which presented some criteria being graded under 3, mainly in PDw FS axial and T2w sagittal FSE sequences. Image quality parameters (noise, contrast, sharpness, blurring, and fluid–soft tissue contrast) also reached higher maximum (up to 3.97) and substantially higher minimum with DRB. Clinical acceptance was therefore clearly better for DRB, corroborated by VGC results. These findings are consistent with prior knee MRI studies showing that deep-learning reconstruction can improve subjective image quality while enabling substantial acceleration without compromising diagnostic performance.[19,58,59]

The T2w sagittal FSE sequence showed the largest performance gap, with DRB scoring up to 3.97 versus 3.07 for standard, likely linked to pixel size differences, where standard used larger voxels, reducing acquisition time but compromising image quality.

In most comparable studies, the number of participants is substantially higher.[43,60] The study by Herrmann et al. included 60 volunteer patients presenting pathology and also demonstrated that the implementation of AI in MR image reconstruction improves image quality and significantly reduces acquisition time, consistent with the findings of the present study. Multiple investigations have demonstrated that deep-learning reconstruction can maintain diagnostic performance, while improve subjective image quality and enabling substantial time reduction, commonly in the 20–50 % range depending on protocol and acceleration strategy.[17,19,61,62] Recht et al. highlighted that a fast DL reconstructed protocol could clinically replace the standard imaging protocol for internal derangement.[61] In

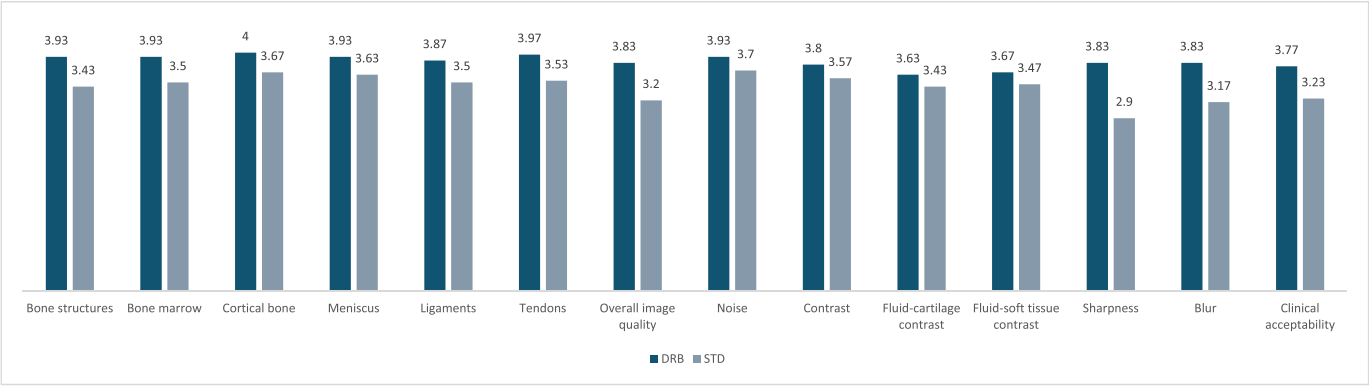


Figure 8. Visual grading analysis of Sagittal T1w FSE.

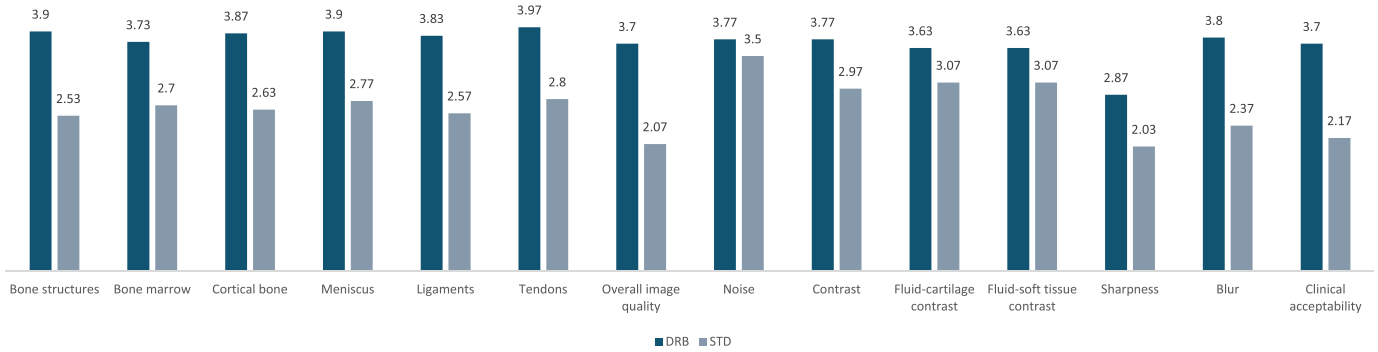


Figure 9. Visual grading analysis of Sagittal T2w FSE.

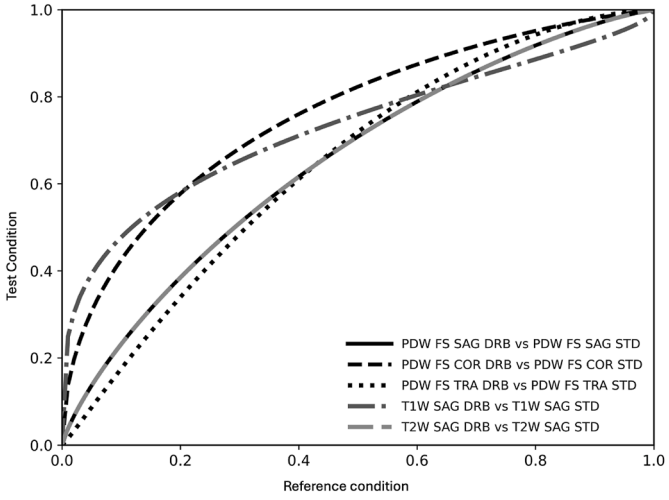


Figure 10. Area under the curve for visual grading characteristics (AUCVGC).

contrast to these studies, the present work is a pilot study, which explains the smaller sample size and focuses on evaluating the impact of acquisition parameters modification, which optimises translation into clinical practice and provides increased understanding for MRI users. This study optimisation focus complements prior larger validation studies by clarifying which acquisition lever most strongly affects time and perceived quality when paired with DL reconstruction.[58,63]

In the current study, several acquisition parameters demonstrated a direct influence on scan duration and image quality. First, reducing spatial resolution in the phase encoding direction effectively shortened acquisition time, since fewer phase-encoding lines were required. The increase in pixel size additionally yielded a higher signal per voxel, thereby improving SNR. While this optimisation is beneficial for low-signal environments, it inherently comes at the expense of spatial detail and should be balanced according to diagnostic requirements. Parallel imaging also contributed to reduced acquisition time without modifying matrix size; however, this gain is counterbalanced by a progressive loss of

Table 5
Area under the curve for visual grading characteristics (AUCVGC) values and interpretation.

Test condition	Reference condition	AUC_{VGC}	95 % Confidence interval	P-value	Interpretation
Sagittal PDw FS DRB	Sagittal PDw FS STD	0.81	0.74–0.85	<0.05	DRB performs statistically better
Coronal PDw FS DRB	Coronal PDw FS STD	0.81	0.77–0.85	<0.05	DRB performs statistically better
Axial PDw FS DRB	Axial PDw FS STD	0.76	0.72–0.80	<0.05	DRB performs statistically better
Sagittal T1w FS DRB	Sagittal T1w FS STD	0.78	0.69–0.86	<0.05	DRB performs statistically better
Sagittal T2w FS DRB	Sagittal T2w FS STD	0.81	0.76–0.85	<0.05	DRB performs statistically better

Table 6
Interobserver agreement results.

	Sagittal PDw FS	Coronal PDw FS	Axial PDw FS	Sagittal T1w	Sagittal T2w	General
AI protocol	0.45	0.46	0.35	0.71	0.65	0.53
Standard protocol	0.15	0.14	0.12	0.71	0.43	0.27
Total agreement	0.29	0.29	0.23	0.52	0.52	0.41

SNR as acceleration factors increase, due to reduced sampled data. Excessive acceleration can further lead to aliasing and phase-related artefacts, highlighting the need for prudent selection of iPAT values aligned with image quality expectations.[40,64] Moreover, the requirement for a reference scan introduces additional complexity, as coil sensitivity mapping must be performed before or during acquisition.[64,65] Many clinical knee studies do not explicitly discuss these parameters, likely because these steps are embedded in the basic protocol. Nonetheless, it remains relevant for protocol reproducibility. The reference scan is often not highlighted in published clinical studies, as other acquisition parameters limit translation and implementation into clinical practice. Similarly, the implementation of FSE sequences enabled substantial time reduction by acquiring multiple echoes within a single repetition time. Nevertheless, longer echo trains have been associated with blurring, phase distortions, and a decrease in overall SNR, again demonstrating a trade-off between speed and image fidelity.[32] Finally, oversampling improved image quality by reducing aliasing and undersampling artefacts, as highlighted in previous studies, though at the cost of slightly longer scan times[19,61,66] ref. Collectively, these considerations emphasise that acceleration strategies must balance acquisition speed with diagnostic image quality. Using AI reconstruction methods allows for acquiring low image quality with reduced acquisition time and enhancing it after acquisition. Future research should specifically address the risk of AI-induced hallucinations and artefacts, as denoising or accelerated reconstruction networks may create anatomically plausible but false structures. While our results mainly showed increased aliasing with low oversampling, as well as a slightly higher presence of artefacts in the final AI-reconstructed sequences, recent work with Deep Resolve Boost reported new artefacts and reduced anatomical delineation, despite preserved lesion visibility. The literature highlights that DL reconstruction can shift artefacts, some become more visible after denoising, and that different commercial tools vary in artefact behaviour, reinforcing the need for structured evaluation.[31,67] Similarly, a recent review[21] highlighted limited scientific evidence regarding diagnostic reliability and artefact behaviour across commercially available AI tools for reconstruction methods. Structured evaluation protocols using blinded comparisons with standard reconstructions and ground-truth validation will therefore be essential in future studies.

Limitation

This study has several limitations. The small sample size limits statistical power but is acceptable for a pilot study. Only healthy knees and a limited number of sequences were evaluated. In addition, absolute VGA has inherent limitations, including the lack of a reference image and potential intra- and inter-observer variability, as well as observer preference effects that may influence scoring.[41]

Future studies with larger sample sizes, pathological cases and multiple observers are needed to better reflect clinical practice. For radiographers and radiologists, these findings highlight the potential of AI-based reconstruction to be integrated into routine workflows. Reduced acquisition times not only improve patient comfort but also allow more patient slots per day, enhancing departmental efficiency. Additionally, the cost-effectiveness of implementing AI software such as DRB should be further explored.[68]

Conclusion

This study demonstrates that the integration of AI-based reconstruction in knee MRI examinations leads to a statistically significant improvement in overall image quality, while simultaneously enabling a substantial reduction in acquisition time of 36.9 %. The application of Deep Resolve Boost (DRB) and Sharp (DRS) across the five selected sequences consistently preserved anatomical fidelity and yielded high-quality images, without compromising the anatomical reproduction of key musculoskeletal components. These findings suggest that AI-based reconstruction techniques can enhance image quality in routine knee MRI protocols, supporting their potential role in optimising clinical workflow and patient comfort, while maintaining diagnostically robust image quality. Future studies should consider including participants with pathology to evaluate the sensitivity and specificity.

Ethics approval and consent to participate

Ethical approval for this study has been granted by the Ethics Committee Vaud (CER-VD, Project-ID: 2024-01679).

Availability of data

Data required for this study may be made available by the author(s) upon reasonable request.

Author contributions

SGH: Writing - Original Draft, Conceptualisation, Methodology, Formal analysis, Writing- Reviewing and Editing.
LBU: Conceptualisation, Methodology, Investigation, Formal analysis, Writing- Reviewing and Editing.
LGA: Conceptualisation, Methodology, Investigation, Formal analysis, Writing- Reviewing and Editing.
YCO: Conceptualisation, Methodology, Investigation, Visualisation, Writing- Reviewing and Editing.
JMcN: Visualisation, Writing- Reviewing and Editing.
AMcG: Visualisation, Writing- Reviewing and Editing.
CSR: Conceptualisation, Methodology, Supervision, Writing-Reviewing and Editing.

Generative AI use

Not applicable.

Funding

The authors have no funding sources to declare.

Conflict of interest statement

JMcN is the current Editor in Chief at Radiography, however, as a co-author of this submission he had no role in or visibility of the handling of the manuscript through the editorial or peer review process.

The remaining authors declare that they have no competing interests.

Acknowledgements

The authors would like to thank the Centre d'Imagerie Diagnostique de Lausanne, for providing access to their MRI machine for data collection.

References

1. Gerena LA, Mabrouk A, DeCastro A. *Knee effusion*. Treasure Island (FL); 2025.
2. Potric A, Mach T, Pereira AC. Miozzari. Gonalgies : Quelle imagerie. *Rev Med Suisse [Internet]*. 2013;9(399):1738–1742. https://www.revmed.ch/revue-medicale-suisse/2013/revue-medicale-suisse-399/gonalgies-quelle-imagerie.
3. Piccolo CL, Mallio CA, Vaccarino F, Grasso RF, Zobel BB. Imaging of knee osteoarthritis: a review of multimodal diagnostic approach. *Quant Imag Med Surg*. 2023;13(11):7582–7595.
4. Chien A, Weaver JS, Kinne E, Omar I. Magnetic resonance imaging of the knee. *Pol J Radiol*. 2020;85:e509–e531.
5. Nacey NC, Geeslin MG, Miller GW, Pierce JL. Magnetic resonance imaging of the knee: an overview and update of conventional and state of the art imaging. *J Magn Reson Imag*. 2017;45(5):1257–1275.
6. Dostál M, Jurasová K, Keřkovský M, Vaníček J, Kalas L, Látal L, et al. What factors affect a patient's subjective perception of MRI examination. *Sci Rep*. 2024;14(1):22731.
7. van Beek EJR, Kuhl C, Anzai Y, Desmond P, Ehman RL, Gong Q, et al. Value of MRI in medicine: more than just another test? *J Magn Reson Imaging [Internet]*. 2019;49(7):e14–e25. https://onlinelibrary.wiley.com/doi/abs/10.1002/jmri.26211.
8. Lawal O, Regelous P, Omiyi D. Supporting claustrophobic patients during magnetic resonance imaging examination– the radiographer perspective. *Radiography [Internet]*. 2024;30(1):80–86. https://www.sciencedirect.com/science/article/pii/S1078817423002067.
9. Sartoretti E, Sartoretti T, Wyss M, van Smoorenburg L, Eichenberger B, van der Duim S, et al. Impact of acoustic noise reduction on patient experience in routine clinical magnetic resonance imaging. *Acad Radiol*. 2022;29(2):269–276.
10. Dean Deyle G. The role of MRI in musculoskeletal practice: a clinical perspective. *J Man Manip Ther*. 2011;19(3):152–161.
11. Ghadimi M, Thomas A. *Magnetic resonance imaging contraindications*. Treasure Island (FL); 2025.
12. Boutin RD, Eshed I, Kassarjian A, Vemuri NV. The global reading room: knee MRI protocols. *AJR Am J Roentgenol*. 2022;219(2):347–348.
13. Havsteen I, Ohlhues A, Madsen KH, Nybing JD, Christensen H, Christensen A. Are movement artifacts in magnetic resonance imaging a real Problem?-A narrative review. *Front Neurol*. 2017;8:232.
14. Stogiannos N, Carlier S, Harvey-Lloyd JM, Brammer A, Nugent B, Cleaver K, et al. A systematic review of person-centred adjustments to facilitate magnetic resonance imaging for autistic patients without the use of sedation or anaesthesia. *Autism*. 2022;26(4):782–797.
15. Budrys T, Veikutis V, Lukosevicius S, Gleizniene R, Monastyreckiene E, Kulakiene I. Artifacts in magnetic resonance imaging: how it can really affect diagnostic image quality and confuse clinical diagnosis? *J Vibroengineering*. 2018;20(2):1202–1213.
16. Chaudhari AS, Fang Z, Kogan F, Wood J, Stevens KJ, Gibbons EK, et al. Super-resolution musculoskeletal MRI using deep learning. *Magn Reson Med*. 2018;80(5):2139–2154.
17. Herrmann J, Koerzdoerfer G, Nickel D, Mostapha M, Nadar M, Gassenmaier S, et al. Feasibility and implementation of a deep learning mr reconstruction for tse sequences in musculoskeletal imaging. *Diagnostics*. 2021;11(8):1–16.
18. Kiryu S, Akai H, Yasaka K, Tajima T, Kunimatsu A, Yoshioka N, et al. Clinical impact of deep learning reconstruction in MRI. *Radiographics*. 2023;43.
19. Johnson PM, Recht MP, Knoll F. Improving the speed of MRI with artificial intelligence. *Semin Muscoskel Radiol*. 2020;24(1):12–20.
20. Herrmann J, Keller G, Gassenmaier S, Nickel D, Koerzdoerfer G, Mostapha M, et al. Feasibility of an accelerated 2D-multi-contrast knee MRI protocol using deep-learning image reconstruction: a prospective intraindividual comparison with a standard MRI protocol. *Eur Radiol*. 2022;32(9):6215–6229.
21. Fransen SJ, Roest C, Simonis FFJ, Yakar D, Kwee TC. The scientific evidence of commercial AI products for MRI acceleration: a systematic review. *Eur Radiol*. 2025;35(8):4736–4746.
22. Foti G, Longo C. Deep learning and AI in reducing magnetic resonance imaging scanning time: advantages and pitfalls in clinical practice. *Pol J Radiol*. 2024;89:e443–e451.
23. Reschke P, Gotta J, Gruenewald LD, Bachir AA, Strecker R, Nickel D, et al. Deep learning in knee MRI: a prospective Study to enhance efficiency, diagnostic confidence and sustainability. *Acad Radiol [Internet]*. 2025;32(6):3585–3596. https://www.sciencedirect.com/science/article/pii/S1076633225002156.
24. Dratsch T, Zäske C, Siedek F, Rauen P, Hokamp NG, Sonnabend K, et al. Reconstruction of 3D knee MRI using deep learning and compressed sensing: a validation study on healthy volunteers. *Eur Radiol Exp [Internet]*. 2024;8(1). https://doi.org/10.1186/s41747-024-00446-0.
25. Wen S, Xu Y, Yang G, Huang F, Zeng Z. Evaluation of deep learning reconstruction in accelerated knee MRI: comparison of visual and diagnostic performance metrics. *Die Radiol [Internet]*. 2025;65(1):74–80. https://doi.org/10.1007/s00117-025-01464-8.
26. Heckel R, Jacob M, Chaudhari A, Perlman O, Shimron E. Deep learning for accelerated and robust MRI reconstruction. *Magn Reson Mater Physics, Biol Med [Internet]*. 2024;37(3):335–368. https://doi.org/10.1007/s10334-024-01173-8.
27. Nikolova E, Kroschke J, Obermüller C, Zecca F, Pawlus K, Jungblut L, et al. Deep-learning reconstructed 3D MRI for comprehensive knee assessment: Comparison with a multisequence 2D protocol at 1.5 T. *Eur J Radiol [Internet]*. 2025;193:112458. https://www.sciencedirect.com/science/article/pii/S0720048X25005443.
28. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. *Nat Rev Cancer [Internet]*. 2018;18(8):500–510. https://www.nature.com/articles/s41568-018-0016-5.
29. Debs P, Fayad LM. The promise and limitations of artificial intelligence in musculoskeletal imaging. *Front Radiol*. 2023;3:1242902.
30. Shimron E, Perlman O. *AI in MRI: computational frameworks for a faster, optimized, and automated imaging workflow*. vol. 10. Basel, Switzerland). Switzerland: Bioengineering; 2023.
31. Hakim A, Rohner R, Winklehner A, Rossel JB, Lehmann C, Wiest R, et al. Deep resolve boost in 2D MRI for neuroradiology: a comparative evaluation of diagnostic gains and potential risks. *Am J Neuroradiol*. 2025:AJNR.A9081.
32. Champendal M, De Labouchère S, Ghotra SS, Gremion I, Sun Z, Torre S, et al. Perspectives of medical imaging professionals about the impact of AI on Swiss radiographers. *J Med Imag Radiat Sci*. 2024;55(4):101741.
33. Hasani N, Morris MA, Rhamim A, Summers RM, Jones E, Siegel E, et al. Trustworthy artificial intelligence in medical imaging. *Pet Clin*. 2022;17(1):1–12.
34. Potočnik J, Foley S, Thomas E. Current and potential applications of artificial intelligence in medical imaging practice: a narrative review. *J Med Imag Radiat Sci*. 2023;54(2):376–385.
35. Hanspach J, Nagel AM, Hensel B, Uder M, Koros L, Laun FB. Sample size estimation: current practice and considerations for original investigations in MRI technical development studies. *Magn Reson Med*. 2021;85(4):2109–2116.
36. Ding C, Cicuttini F, Scott F, Cooley H, Jones G. Association between age and knee structural change: a cross sectional MRI based study. *Ann Rheum Dis*. 2005;64(4):549–555.
37. In J. Introduction of a pilot study. *Korean J Anesthesiol*. 2017;70(6):601–605.
38. Behl N, Ph D. *Deep resolve : unrivaled speed in MRI*. vol. 89. 2024:2–11.
39. Nicolas Behl. Deep resolve – mobilizing the power of networks. *MAGNETOM Flash*. 2021;78(78).
40. Vosshenrich J, Koerzdoerfer G, Fritz J. Modern acceleration in musculoskeletal MRI: applications, implications, and challenges. *Skelet Radiol*. 2024;53(9):1799–1813.
41. Precht H, Hansson J, Outzen C, Hogg P, Tingberg A. Radiographers' perspectives' on visual grading analysis as a scientific method to evaluate image quality. *Radiography [Internet]*. 2019;25:S14–S18. https://doi.org/10.1016/j.radi.2019.06.006.
42. Lopez Schmidt I, Haag N, Shahzadi I, Frohwein LJ, Schneider C, Niehoff JH, et al. Diagnostic image quality of a low-field (0.55T) knee MRI protocol using deep learning image reconstruction compared with a standard (1.5T) knee MRI protocol. *J Clin Med*. 2023;12(5):1–12.
43. Herrmann J, Keller G, Gassenmaier S, Nickel D, Koerzdoerfer G, Mostapha M, et al. Feasibility of an accelerated 2D-multi-contrast knee MRI protocol using deep-learning image reconstruction: a prospective intraindividual comparison with a standard MRI protocol. *Eur Radiol*. 2022;32(9):6215–6229.
44. Attard S, Castillo J, Zarb F. Establishment of image quality for MRI of the knee joint using a list of anatomical criteria. *Radiography*. 2018;24(3):196–203.
45. Cheng Q, Zhao FC. Comparison of 1.5- and 3.0-T magnetic resonance imaging for evaluating lesions of the knee: a systematic review and meta-analysis (PRISMA-compliant article). *Med (United States)*. 2018;97(38):1–9.
46. Svalkvist A, Svensson S, Hagberg T, Båth M. Viewdex 3.0 - recent development of a software application facilitating assessment of image quality and observer performance. *Radiat Protect Dosim*. 2021;195(3–4):372–377.
47. Börjesson S, Håkansson M, Båth M, Kheddache S, Svensson S, Tingberg A, et al. A software tool for increased efficiency in observer performance studies in radiology. *Radiat Protect Dosim*. 2005;114(1–3):45–52.
48. Svensson S, Zachrisson S, Svalkvist A, Ba M, Ha M, Radiology D, et al. V iew dex : an efficient and easy-to-use softw are. *Radiat Protect Dosim*. 2010;139(1):42–51.

49. Bath M. *VIEWDEX : a status report*. vol. 169. 2016:38–45 (1).
50. Båth M, Månsson LG. Visual grading characteristics (VGC) analysis: a non-parametric rank-invariant statistical method for image quality evaluation. *Br J Radiol*. 2007;80(951):169–176.
51. Bath M. *Evaluating imaging systems : practical applications*. vol. 139. 2010: 26–36 (1).
52. Bath M, Hansson J. Vgc analyzer : a SoftwAre for statistical analysis of fully crossed multiple-reader multiple-case visual grading characteristics studies. *Radiat Protect Dosim*. 2016:1–8.
53. Hansson J, Månsson LG, Båth M. The validity of using ROC software for analysing visual grading characteristics data: an investigation based on the novel software VGC analyzer. *Radiat Protect Dosim*. 2016;169(1): 54–59.
54. Hansson J, Månsson LG, Båth M. Evaluation of VGC analyzer by comparison with gold standard ROC software and analysis of simulated visual grading data. *Radiat Protect Dosim*. 2021;195(3–4):378–390.
55. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159.
56. Benchoufi M, Matzner-Lober E, Molinari N, Jannot AS, Soyer P. Interobserver agreement issues in radiology. *Diagn Interv Imaging [Internet]*. 2020;101(10): 639–641. https://doi.org/10.1016/j.diii.2020.09.001.
57. Chen G, Faris P, Hemmelgarn B, Walker RL, Quan H. Measuring agreement of administrative data with chart data using prevalence unadjusted and adjusted kappa. *BMC Med Res Methodol*. 2009;9(1):1–8.
58. Herrmann J, Keller G, Gassenmaier S, Nickel D, Koerzdoerfer G, Mostapha M, et al. Feasibility of an accelerated 2D-multi-contrast knee MRI protocol using deep-learning image reconstruction: a prospective intra-individual comparison with a standard MRI protocol. *Eur Radiol [Internet]*. 2022;32(9):6215–6229. https://doi.org/10.1007/s00330-022-08753-z.
59. Lee J, Jung M, Park J, Kim S, Im Y, Lee N, et al. Highly accelerated knee magnetic resonance imaging using deep neural network (DNN)–based reconstruction: prospective, multi-reader, multi-vendor study. *Sci Rep [Internet]*. 2023;13(1): 17264. https://doi.org/10.1038/s41598-023-44248-7.
60. Hahn S, Yi J, Lee HJ, Lee Y, Lim YJ, Bang JY, et al. Image quality and diagnostic performance of accelerated shoulder MRI with deep learning-based reconstruction. *AJR Am J Roentgenol*. 2022;218(3):506–516.
61. Recht MP, Zbontar J, Sodickson DK, Knoll F, Yakubova N, Sriram A, et al. Using deep learning to accelerate knee MRI at 3 T: results of an interchangeability study. *AJR Am J Roentgenol*. 2020;215(6):1421–1429.
62. Kim M, Lee SM, Park C, Lee D, Kim KS, Jeong HS, et al. Deep learning-enhanced parallel imaging and simultaneous multislice acceleration reconstruction in knee MRI. *Investig Radiol*. 2022;57(12):826–833.
63. Lee SM, Kim M, Park C, Lee D, Kim KS, Jeong HS, et al. Deep learning-reconstructed parallel accelerated imaging for knee MRI. *Curr Med imaging*. 2024;20:e240523217293.
64. Deshmane A, Gulani V, Griswold MA, Seiberlich N. Parallel MR imaging. *J Magn Reson Imag*. 2012;36(1):55–72.
65. Thompson T. *Expert insights : hidden gems from application specialists at Siemens Healthineers*. vol. 89. 2024:2–5.
66. Huang SY, Seethamraju RT, Patel P, Hahn PF, Kirsch JE, Guimaraes AR. Body MR imaging: artifacts, k-Space, and solutions. *Radiogr a Rev Publ Radiol Soc North Am Inc*. 2015;35(5):1439–1460.
67. Kiryu S, Akai H, Yasaka K, Tajima T, Kunimatsu A, Yoshioka N, et al. Clinical impact of deep learning reconstruction in MRI. *RadioGraphics [Internet]*. 2023;43(6):e220133. https://doi.org/10.1148/rg.220133.
68. Brix MAK, Järvinen J, Bode MK, Nevalainen M, Nikki M, Niinimäki J, et al. Financial impact of incorporating deep learning reconstruction into magnetic resonance imaging routine. *Eur J Radiol*. 2024;175:111434.