

Article

Accuracy of Deep Learning-Driven MR Arthrography of the Shoulder: Compressed 3D in Comparison to Standard FSE Sequences

Gianluca Tripodi ¹, Flavio Spoto ^{2,*}, Giuseppe Ocello ¹, Leonardo Monterubbiano ², Paolo Avanzi ³ and Giovanni Foti ²

¹ Department of Radiology, Policlinico G. Martino, 98124 Messina, Italy

² Department of Radiology, IRCCS Sacro Cuore Hospital, 37024 Negrar, Italy; giovanni.foti@sacrocuore.it (G.F.)

³ Department of Orthopedic Surgery, IRCCS Sacro Cuore Hospital, 37024 Negrar, Italy

* Correspondence: flaviospoto@hotmail.it

Abstract

Background/Objectives: Magnetic resonance arthrography is the reference standard for evaluating glenoid labral lesions. Deep learning (DL) reconstruction algorithms may accelerate 3D acquisitions while maintaining image quality. This study assesses the diagnostic accuracy of DL-based isotropic 3D MR imaging for detecting glenoid labral lesions. **Methods:** This prospective study included 128 consecutive patients (79 men, 49 women; mean age 38.4 years) undergoing shoulder MR arthrography between June 2023 and April 2025. DL-based 3D sequences (acquisition time: 3:26) were compared with conventional multiplanar TSE and PD-FS sequences (acquisition time: 24–28 min). Two independent radiologists assessed glenoid labral lesions, bone marrow edema, and rotator cuff abnormalities using a four-point Likert scale. Sensitivity, specificity, and interobserver agreement were calculated. **Results:** DL-based 3D sequences demonstrated 94.7–95.1% sensitivity and 100% specificity for glenoid labral lesions, with excellent interobserver agreement ($\kappa = 0.812$). The area under the ROC curve was 0.894. Combined 3D protocols (T1 + PD-FS) showed superior accuracy (97.8%) compared to single sequences (90.5%, $p = 0.012$). For bone marrow edema, sensitivity was 82.9% with 100% specificity. Rotator cuff evaluation achieved 75% sensitivity with 100% specificity. **Conclusions:** DL-based isotropic 3D sequences provide high diagnostic accuracy for glenoid labral pathology while reducing scan time by 75%. Combined T1 and PD-FS protocols optimize performance. These findings support selective implementation of DL-accelerated 3D protocols in shoulder MR arthrography, particularly for labral assessment, while acknowledging that conventional protocols may remain preferable in specific clinical scenarios.

Keywords: shoulder; MRI arthrography; 3D; deep learning; artificial intelligence; labral tear



Academic Editor: Ralf Bieger

Received: 28 November 2025

Revised: 1 January 2026

Accepted: 10 February 2026

Published: 27 February 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Magnetic resonance arthrography (MRA) of the shoulder is the non-invasive gold standard for assessing intra-articular disorders, particularly glenoid labral lesions associated with glenohumeral instability [1–6]. The technique involves direct injection of diluted gadolinium into the joint space, distending the capsule and enhancing visualization of labral tears, cartilage defects, and ligamentous abnormalities. Conventional multiplanar turbo spin-echo (TSE) sequences provide excellent sensitivity for labral pathology but are limited by prolonged acquisition times (20–30 min), motion artifact susceptibility, and restricted anatomical coverage [7–9]. These limitations may become critical when

detecting subtle sub-millimetric labral abnormalities, while extended scan times reduce patient comfort and scanner throughput. Recent advances in deep learning (DL) algorithms have enabled high-quality image reconstruction from undersampled datasets, facilitating substantial scan time reductions [10–16]. These algorithms utilize convolutional neural networks to map undersampled k-space data to fully sampled images, effectively recovering missing information while suppressing aliasing artifacts. In musculoskeletal imaging, DL reconstruction has shown promise in maintaining diagnostic quality despite acceleration. Reschke et al. [17] demonstrated that DL-based knee MRI sequences achieved up to 4-fold acceleration while preserving diagnostic performance for meniscal and cartilage pathology. Three-dimensional imaging offers theoretical advantages for shoulder arthrography, including isotropic resolution enabling multiplanar reconstructions without quality loss, complete joint coverage eliminating inter-slice gaps, and improved detection of obliquely oriented tears [18–20]. However, 3D sequences traditionally require longer acquisition times than 2D protocols. DL acceleration could overcome this limitation by enabling rapid 3D acquisitions with submillimeter isotropic resolution [18–20]. Despite these potential benefits, no studies have systematically evaluated DL-reconstructed isotropic 3D sequences for labral pathology assessment in shoulder MRA. Furthermore, accelerated 3D acquisitions may compromise tissue contrast compared to standard TSE sequences, potentially affecting diagnostic accuracy for subtle intra-articular pathology where contrast enhancement is critical. The primary aim was to compare the diagnostic performance of DL-accelerated isotropic 3D MRA sequences against standard 3 mm TSE and PD-FS protocols for detecting glenoid labral lesions. Secondary objectives included assessing rotator cuff abnormalities, bone marrow edema detection, inter-observer agreement, and subgroup analysis based on available sequence combinations.

2. Materials and Methods

2.1. Study Population

This prospective study was conducted at a tertiary referral institution between June 2023 and April 2025. Of the 133 consecutive patients referred for shoulder MR arthrography, 128 were included after excluding 4 with incomplete datasets and 1 with severe metal artifacts. Inclusion criteria comprised a clinical indication for shoulder arthrography and acquisition of at least one DL-accelerated 3D sequence with complete standard 2D sequences. The institutional review board approved the study (Prot. n.33264; Prog. 4287CESC), and all patients provided written informed consent.

2.2. Imaging Protocol

All patients underwent fluoroscopy-guided intra-articular injection of 12–15 mL diluted gadolinium (Gd-DTPA, 1–2 mmol/L) using an anterior approach. MR imaging was performed within 10 min on a 3.0 T system (uMR Omega, United Imaging Healthcare Shanghai, Shanghai, China) with a 16-channel shoulder coil.

Standard 2D protocol included:

- TSE T1-weighted: TR/TE = 650/18 ms, 3 mm slices, acquired in axial, coronal oblique, and radial planes (3:42 per plane)
- PD-FS: TR/TE = 4800/46 ms, 3 mm slices, acquired in axial and coronal oblique planes (4:52 per plane)
- Total acquisition time: 24–28 min

DL-accelerated 3D sequences using Accelerated Compressed Sensing (ACS) with proprietary deep learning reconstruction (uAI):

- 3D TSE T1: TR/TE = 450/17.58 ms, 0.8 mm isotropic voxels, FOV = 160 × 160 × 120 mm
- 3D PD-FS: TR/TE = 1500/30 ms, 0.84 mm isotropic voxels, FOV = 160 × 160 × 120 mm

- Acceleration factor: 2.5×, acquisition time: 3:26 per sequence

The DL reconstruction algorithm employs a multiscale convolutional neural network trained on paired full/undersampled datasets, providing real-time reconstruction (<1 min) through iterative k-space consistency and image domain refinement.

2.3. Image Analysis

Two musculoskeletal radiologists (15 and 4 years of experience) independently evaluated all studies, blinded to clinical data. Standard 2D sequences served as the reference standard. It should be noted that the reported sensitivity and specificity values reflect agreement with standard 2D MRA rather than absolute diagnostic accuracy, as arthroscopic correlation was not available. After ≥ 1 month washout, readers assessed 3D sequences using multiplanar reconstructions. Findings were scored 1–4 (definitely absent to definitely present). Readers were aware of the study design, including that 2D sequences would serve as the reference standard. However, several measures were implemented to minimize bias: (1) 3D and 2D interpretations were performed in separate sessions with a minimum 4-week washout period to reduce recall bias; (2) readers documented their findings using a standardized structured reporting template before any comparison; (3) cases were presented in randomized order for each reading session; and (4) readers were instructed to interpret each protocol independently based solely on the imaging findings without reference to their prior assessments.

The glenoid labrum was evaluated in three regions:

- Superior: SLAP lesions
- Anteroinferior: Bankart variants (including ALPSA, GLAD)
- Posterior: including POLPSA

Labral tears were diagnosed by a linear hyperintense signal, irregular morphology, or displacement. Bone marrow edema required increased PD-FS signal or corresponding T1 hypointensity. Rotator cuff assessment included tendinopathy, partial, and full-thickness tears. Each reader independently evaluated both 2D (reference) and 3D sequences, with diagnostic performance calculated against their own 2D assessments. This design evaluates within-reader consistency between modalities rather than agreement with a single consensus reference, providing a clinically relevant assessment of each reader's ability to maintain diagnostic consistency when transitioning to accelerated 3D protocols.

2.4. Subgroup Analysis

Patients were stratified by available 3D sequences:

- Group A ($n = 46$): Both T1 and PD-FS available
- Group B ($n = 40$): T1 only
- Group C ($n = 42$): PD-FS only

The distribution of 3D sequences across patients resulted from practical constraints during the study period. Group A ($n = 46$) received both 3D T1 and PD-FS sequences when scanner time permitted complete protocol acquisition. Group B ($n = 40$, 3D T1 only) and Group C ($n = 42$, 3D PD-FS only) resulted from time limitations when patient scheduling required abbreviated examinations, technical factors necessitating sequence repetition that precluded completion of both 3D acquisitions, or early protocol development phases when sequence parameters were being optimized. This pragmatic approach reflects real-world clinical implementation where complete ideal protocols are not always achievable, and the resulting subgroup analysis provides clinically relevant insights into single-sequence performance when time constraints mandate protocol abbreviation.

2.5. Statistical Analysis

Diagnostic performance metrics were calculated with 95% confidence intervals using the Wilson method. ROC analysis determined optimal thresholds via Youden's index. Inter-observer agreement used weighted Cohen's kappa. Subgroup comparisons employed chi-square tests with Bonferroni correction ($\alpha = 0.017$). Sample size calculation indicated 124 patients needed for 90% expected sensitivity with 5% margin of error. Analyses were performed using SPSS 28.0 and R 4.3.2.

3. Results

3.1. Study Population Characteristics

The initial cohort comprised 133 consecutive patients referred for shoulder MR arthrography. Following application of exclusion criteria, 128 patients (79 males, 49 females) were included in the final analysis. The mean age was 38.4 years (range 17–73 years, SD $\pm$ 12.3 years). Clinical indications for arthrography included chronic shoulder pain ($n = 67$, 52.3%), suspected labral tear following trauma ($n = 41$, 32.0%), and recurrent instability ($n = 20$, 15.6%).

3.2. Reference Standard Findings

Analysis of standard 2D MRA sequences revealed a high prevalence of pathology in this referred population. Glenoid labral lesions were identified in 98 of 128 cases (76.6%), distributed as follows: isolated superior labral (SLAP) lesions in 34 cases (34.7%), anteroinferior labral tears (Bankart and variants) in 38 cases (38.8%), posterior labral tears in 8 cases (8.2%), and combined/complex tears in 18 cases (18.4%). The remaining 30 patients (23.4%) showed no significant labral signal alterations. Rotator cuff evaluation revealed abnormalities in 43 of 128 cases (33.6%), including full-thickness tears ($n = 12$, 9.4%), partial-thickness tears ($n = 19$, 14.8%), and tendinopathy without a discrete tear ($n = 12$, 9.4%). The supraspinatus tendon was most commonly affected ($n = 31$, 72.1% of rotator cuff pathology), followed by subscapularis ($n = 8$, 18.6%) and infraspinatus ($n = 4$, 9.3%). Bone marrow edema was present in 50 patients (39.1%), with the following distribution: Hill-Sachs lesions with associated edema ($n = 22$, 44.0%), glenoid rim edema associated with labral pathology ($n = 18$, 36.0%), and greater tuberosity edema ($n = 10$, 20.0%).

3.3. Diagnostic Performance of 3D ACS Sequences

The DL-accelerated 3D sequences demonstrated excellent diagnostic performance for glenoid labral lesion detection (Table 1). In the overall population, Reader 1 achieved sensitivity of 94.7% (89/94, 95% CI: 88.1–97.8%) and specificity of 100% (34/34, 95% CI: 89.8–100%), yielding an accuracy of 96.1% (123/128, 95% CI: 91.1–98.5%). Reader 2 showed comparable performance with sensitivity of 95.1% (78/82, 95% CI: 88.0–98.4%) and specificity of 100% (46/46, 95% CI: 92.3–100%), resulting in accuracy of 96.9% (124/128, 95% CI: 92.2–98.9%). The difference in lesion identification between readers (94 vs. 82 labral lesions on 2D reference) reflects expected inter-reader variability for subtle labral pathology. Importantly, both readers demonstrated excellent internal consistency between their respective 2D and 3D assessments. Inter-reader agreement for 3D interpretation was excellent ($\kappa = 0.812$). ROC curve analysis confirmed the excellent discriminative ability of 3D ACS sequences with an AUC of 0.894 (95% CI: 0.842–0.946, $p < 0.001$). The optimal diagnostic threshold identified by Youden's index was a Likert score ≥ 3 , which corresponded to "probably present" or a higher confidence level (Figure 1). For bone marrow edema detection, 3D sequences achieved an overall sensitivity of 82.9% (34/41, 95% CI: 68.7–91.5%) with perfect specificity of 100% (87/87, 95% CI: 95.8–100%). The slightly lower sensitivity for edema detection likely reflects the inherent contrast characteristics of accelerated 3D sequences compared to dedicated fluid-sensitive 2D imaging. Rotator cuff assessment

revealed limitations of 3D sequences for subtle tendon pathology. Overall sensitivity was 75.0% (21/28, 95% CI: 56.6–87.3%) for Reader 1 and 63.0% (17/27, 95% CI: 44.2–78.5%) for Reader 2, though specificity remained at 100% (95% CI: 96.3–100%) for both readers. The lower sensitivity was primarily attributable to missed small partial-thickness tears and subtle tendinopathy. Accuracy results are displayed in Figures 1 and 2.

Table 1. Diagnostic performance of 3D ACS sequences in the total population ($n = 128$).

Parameter	Reader 1	Reader 2
Glenoid labrum		
Sensitivity	94.7% (89/94, 95% CI: 88.1–97.8%)	95.1% (78/82, 95% CI: 88.0–98.4%)
Specificity	100% (34/34, 95% CI: 89.8–100%)	100% (46/46, 95% CI: 92.3–100%)
PPV	100% (89/89)	100% (78/78)
NPV	87.2% (34/39, 95% CI: 73.3–94.4%)	92.0% (46/50, 95% CI: 81.2–96.8%)
Accuracy	96.1% (123/128, 95% CI: 91.1–98.5%)	96.9% (124/128, 95% CI: 92.2–98.9%)
Bone marrow edema		
Sensitivity	82.9% (34/41, 68.7–91.5%)	92.3% (24/26, 75.9–97.9%)
Specificity	100% (87/87, 95.8–100%)	100% (102/102, 96.4–100%)
PPV	100% (34/34)	100% (24/24)
NPV	92.6% (87/94, 85.5–96.4%)	98.1% (102/104, 93.3–99.5%)
Accuracy	94.5% (121/128, 89.1–97.4%)	98.4% (126/128, 94.5–99.6%)
Rotator cuff		
Sensitivity	75.0% (21/28, 56.6–87.3%)	63.0% (17/27, 44.2–78.5%)
Specificity	100% (100/100, 96.3–100%)	100% (101/101, 96.4–100%)
PPV	100% (21/21)	100% (17/17)
NPV	93.5% (100/107, 87.2–96.8%)	91.0% (101/111, 84.2–95.1%)
Accuracy	94.5% (121/128, 89.1–97.4%)	92.2% (118/128, 86.2–95.8%)

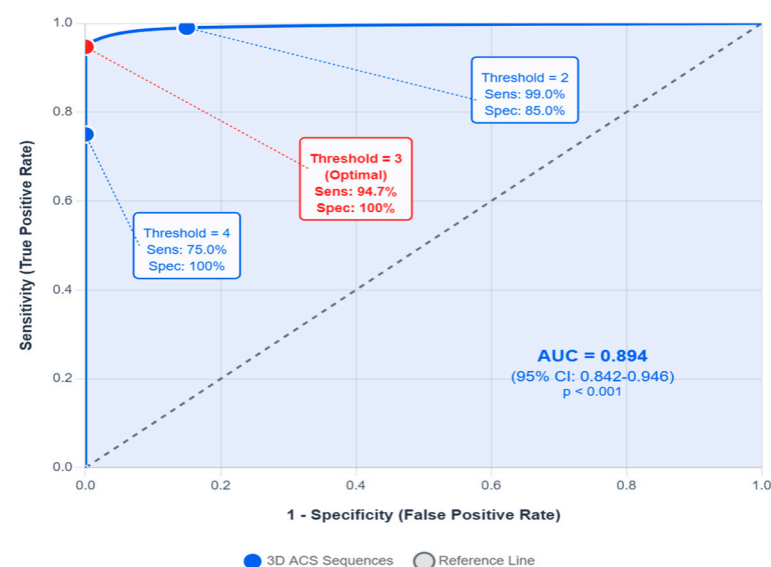


Figure 1. Receiver Operating Characteristic (ROC) curve for detection of glenoid labral lesions using deep learning-accelerated 3D MR arthrography in 128 patients. The area under the curve (AUC) was 0.894 (95% CI: 0.842–0.946, $p < 0.001$), indicating excellent discriminative ability. The optimal diagnostic threshold identified by Youden’s index was a Likert score ≥ 3 (“probably present” or higher confidence level), which yielded 94.7–95.1% sensitivity and 100% specificity (red box). Additional threshold values are shown for comparison to illustrate the trade-off between sensitivity and specificity. The dashed diagonal line represents the reference line of no discrimination (AUC = 0.5).

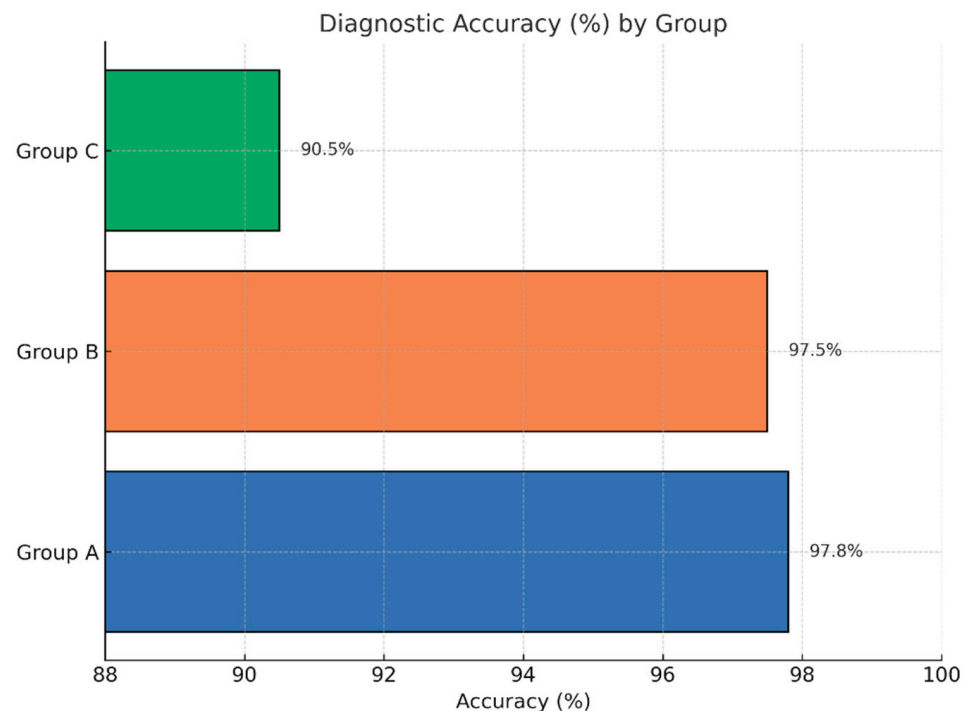


Figure 2. Comparison of diagnostic accuracy for labral lesion detection among protocol subgroups. Group A ($n = 46$) received both 3D T1-weighted and PD-FS sequences; Group B ($n = 40$) received 3D T1-weighted only; Group C ($n = 42$) received 3D PD-FS only. Error bars represent 95% confidence intervals. Group A demonstrated significantly higher accuracy (97.8%) compared to Group C (90.5%, $p = 0.012$ after Bonferroni correction). The difference between Groups A and B was not statistically significant ($p = 0.752$).

3.4. Subgroup Analysis by Protocol Type

Comparative analysis across the three protocol groups revealed statistically significant differences in diagnostic performance ($\chi^2 = 8.74$, $p = 0.013$). Group A, with access to both 3D sequences (T1 and PD-FS), demonstrated superior performance with a sensitivity of 96.8% (30/31, 95% CI: 83.8–99.4%) and accuracy of 97.8% (45/46, 95% CI: 88.7–99.6%). This represented the highest diagnostic yield among all groups (Table 2). Group B (3D T1 only) maintained high performance with sensitivity of 95.0% (19/20, 95% CI: 76.4–99.1%) and accuracy of 97.5% (39/40, 95% CI: 87.1–99.6%). Group C (3D PD-FS only) showed significantly lower performance with a sensitivity of 87.5% (28/32, 95% CI: 71.9–95.2%) and accuracy of 90.5% (38/42, 95% CI: 77.9–96.2%). The difference in accuracy between Groups A and C was statistically significant (difference 7.3%, 95% CI: 1.2–13.4%, $p = 0.012$ after Bonferroni correction). For secondary endpoints, Group A demonstrated a trend toward superior performance in bone marrow edema detection with a sensitivity of 88.9% (16/18, 95% CI: 67.2–96.9%) compared to 78.3% (18/23, 95% CI: 58.1–90.3%) for Group C, though this difference did not reach statistical significance ($p = 0.382$). Interestingly, for rotator cuff evaluation, Group C showed paradoxically higher sensitivity of 83.3% (10/12, 95% CI: 55.2–95.3%) compared to Group A at 63.6% (7/11, 95% CI: 35.4–84.8%), suggesting that PD-FS weighting may be particularly valuable for tendon assessment (Table 3).

Table 2. Comparison of diagnostic performance for labral lesion detection across protocol subgroups. Group A received both 3D T1-weighted and PD-FS sequences; Group B received 3D T1-weighted only; Group C received 3D PD-FS only. Statistical comparison performed using chi-square test with Bonferroni correction for multiple comparisons (adjusted significance threshold $\alpha = 0.017$). p -values represent comparison to Group A (reference). 95% confidence intervals calculated using the Wilson method.

Group	Sensitivity (95% CI)	Specificity (95% CI)	Accuracy (95% CI)	p -Value	n
Group A (T1 + PD-FS)	96.8% (30/31, 83.8–99.4%)	100% (15/15, 79.6–100%)	97.8% (45/46, 88.7–99.6%)	Ref.	46
Group B (T1 only)	95.0% (19/20, 76.4–99.1%)	100% (20/20, 83.9–100%)	97.5% (39/40, 87.1–99.6%)	0.752	40
Group C (PD-FS only)	87.5% (28/32, 71.9–95.2%)	100% (10/10, 72.2–100%)	90.5% (38/42, 77.9–96.2%)	0.012	42

Between-group accuracy differences with 95% confidence intervals: Group A vs. Group B: 0.3% (95% CI: −6.8% to 7.4%); Group A vs. Group C: 7.3% (95% CI: 1.2% to 13.4%); Group B vs. Group C: 7.0% (95% CI: 0.4% to 13.6%).

Table 3. Diagnostic performance for bone marrow edema and rotator cuff abnormalities in Groups A and C. Statistical comparison performed using Fisher’s exact test due to small cell counts. Group B was excluded from this analysis due to the absence of fluid-sensitive (PD-FS) sequences, limiting bone marrow edema assessment. Values represent percentage (n/N , 95% CI) calculated using the Wilson method.

Parameter	Group A ($n = 46$)	Group C ($n = 42$)	p -Value
Bone Marrow Edema			
Sensitivity	88.9% (16/18, 67.2–96.9%)	78.3% (18/23, 58.1–90.3%)	0.382
Specificity	100% (28/28, 87.9–100%)	100% (19/19, 83.2–100%)	1.000
Accuracy	95.7% (44/46, 85.5–98.8%)	88.1% (37/42, 74.7–95.0%)	0.248
Rotator Cuff			
Sensitivity	63.6% (7/11, 35.4–84.8%)	83.3% (10/12, 55.2–95.3%)	0.371
Specificity	100% (35/35, 90.1–100%)	100% (30/30, 88.6–100%)	1.000
Accuracy	91.3% (42/46, 79.7–96.6%)	95.2% (40/42, 84.2–98.7%)	0.677

3.5. Inter-Observer Agreement

Analysis of inter-observer agreement demonstrated excellent concordance between readers for all evaluated structures. For glenoid labral assessment, weighted Cohen’s kappa was 0.834 (95% CI: 0.754–0.914) using standard sequences and 0.812 (95% CI: 0.728–0.896) with 3D ACS sequences, both indicating excellent agreement. The minimal decrease in agreement with 3D sequences suggests that these acquisitions provide consistent diagnostic information across readers with different experience levels. Bone marrow edema evaluation showed the highest inter-observer agreement with $\kappa = 0.887$ (95% CI: 0.812–0.962) for standard sequences and $\kappa = 0.865$ (95% CI: 0.785–0.945) for 3D sequences. Rotator cuff assessment demonstrated $\kappa = 0.845$ (95% CI: 0.759–0.931) with standard sequences and $\kappa = 0.798$ (95% CI: 0.701–0.895) with 3D sequences, the latter at the boundary between substantial and excellent agreement. Overall diagnostic concordance, assessed by Kendall’s coefficient of concordance (W), was 0.912 ($p < 0.001$), indicating excellent agreement between readers regardless of protocol type. McNemar’s test revealed no significant differences in diagnostic performance between readers of different experience levels ($p = 0.453$), suggesting that 3D ACS sequences can be reliably interpreted by radiologists with varying expertise. Agreement results are shown in Figure 3. Figures 4–6 show three explicative cases.

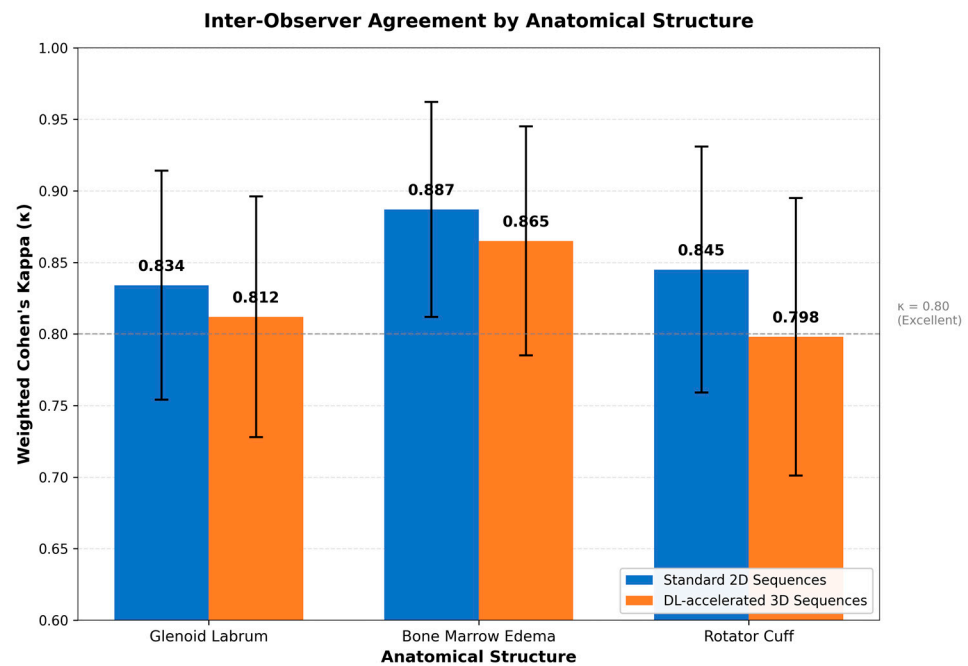


Figure 3. Inter-observer agreement (weighted Cohen's kappa) by anatomical structure for standard 2D sequences (blue bars) and DL-accelerated 3D sequences (orange bars). Error bars represent 95% confidence intervals. All κ values exceeded 0.79, indicating substantial to excellent agreement. The highest agreement was observed for bone marrow edema assessment ($\kappa = 0.887$ for 2D, $\kappa = 0.865$ for 3D).

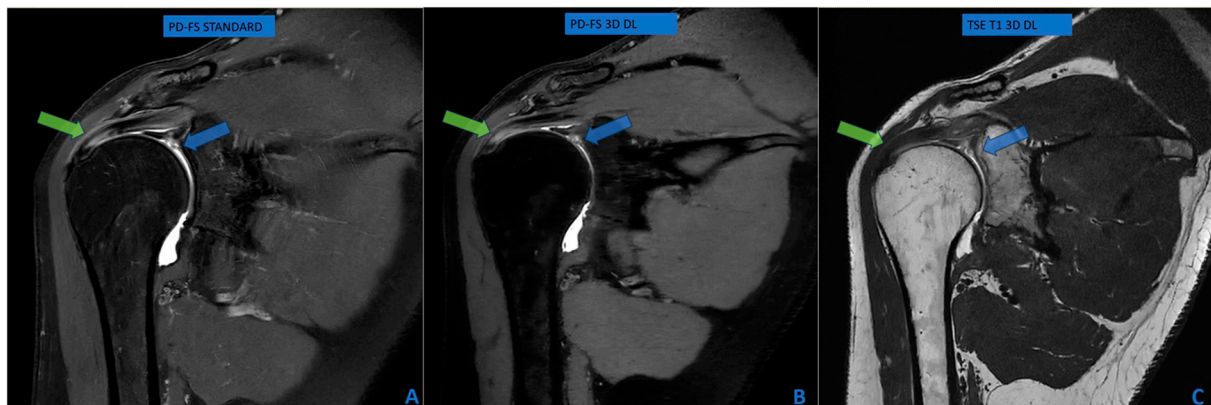


Figure 4. SLAP lesion with marked labral fissuring and supraspinatus partial tendon tear in a 33-year-old man. Coronal MR arthrography images. (A) Standard 3 mm coronal PD-FS sequence clearly demonstrates a small contrast-filled cleft within the superior glenoid labrum (blue arrow). The finding is also well visualized on (B) PD-FS 3D deep learning-accelerated sequence and (C) TSE T1-weighted 3D deep learning-accelerated sequence, both showing excellent contrast resolution between intra-articular contrast and the fissured labrum. A partial-thickness articular-sided tear at the supraspinatus tendon insertion with advanced tendinopathy (green arrow) is evident as a contrast-filled gap on the standard PD-FS image (A), although this finding shows relatively lower contrast resolution on the PD-FS 3D DL and could be missed on TSE T1 3D DL sequences (B,C). This case illustrates the superior labral detection capability of 3D sequences while highlighting their limitation for subtle rotator cuff pathology.

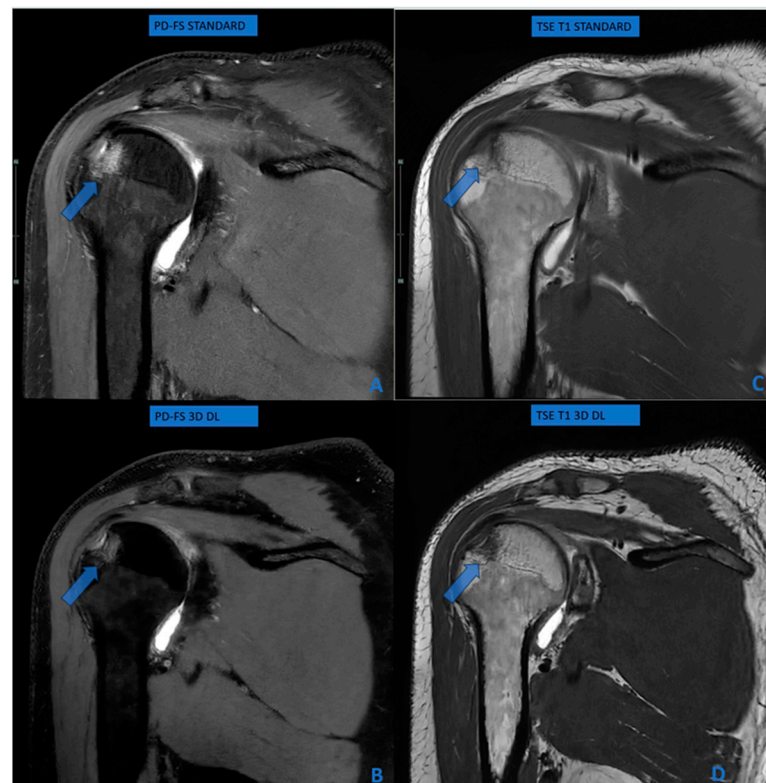


Figure 5. Deep cortical indentation at the posterosuperior humeral head in coronal view, consistent with a Hill-Sachs lesion, associated with prominent cortico-subcortical bone marrow edema (blue arrow). Comparative coronal images from a patient with glenohumeral instability. Bone marrow edema appears as an intramedullary hyperintense area on (A) standard PD-FS sequence and is equally well visualized on (B) PD-FS 3D deep learning–accelerated sequence. Corresponding hypointense signal is seen on (C) standard TSE T1-weighted and (D) TSE T1-weighted 3D deep learning–accelerated sequences. The high spatial and contrast resolution between the bone tissue and the adjacent soft tissue structures of the 3D DL sequences (B,D) improve clearly delineation of the cortical indentation. The isotropic voxel size (0.8 mm) enables high-quality multiplanar reconstructions without step-ladder artifacts.

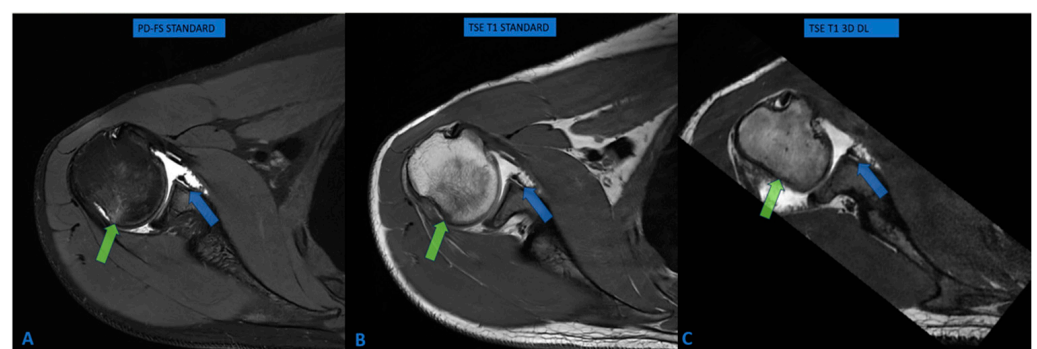


Figure 6. Large cortical indentation at the posterosuperior humeral head with surrounding bone marrow edema, consistent with a Hill-Sachs lesion (green arrow). An associated anterior glenoid labral injury (ALPSA lesion) characterized by detachment of the anterior labrum with intact but stretched and displaced capsuloperiosteal complex (blue arrow) is shown on (A–C). Axial MR arthrography images demonstrating excellent visualization of the ALPSA lesion across all sequences, with the 3D deep learning–accelerated sequences (B,C) providing comparable diagnostic information to the standard 2D protocol (A).

3.6. Analysis of Diagnostic Confidence

Reader confidence scores, based on the 4-point Likert scale, were analyzed to assess the subjective interpretability of 3D sequences. For definite diagnoses (scores 1 or 4), readers assigned confident scores in 78.3% of cases with 3D sequences compared to 85.7% with standard sequences. The difference was most pronounced for subtle pathology, where intermediate confidence scores (2 or 3) were more frequent with 3D imaging.

3.7. Time Efficiency Analysis

The implementation of 3D ACS sequences resulted in substantial time savings. Standard protocol acquisition required 24–28 min, while a complete 3D protocol (both T1 and PD-FS) required only 6 min 52 s, representing a 75.5% reduction in scan time. Even accounting for post-processing and reconstruction time, the total time from acquisition start to image availability was reduced by approximately 70%.

4. Discussion

This prospective study demonstrates that deep learning-accelerated 3D sequences significantly reduce acquisition time in shoulder MR arthrography while maintaining high diagnostic accuracy for glenoid labral pathology. The principal finding of 94.7–95.1% sensitivity with perfect specificity validates the clinical applicability of these advanced reconstruction techniques. Our findings confirm the excellent diagnostic performance of AI-accelerated sequences in assessing RC tears, BM lesion and glenoid labral pathology during shoulder MRI [9,15,21]. Specifically, both readers achieved high sensitivity and specificity and a large area under the ROC curve (AUC 0.894); these findings are in line with those of a recently published paper focused on the shoulder MRI [15]. In the above-mentioned paper, as concerns the rotator cuff evaluation, the 2-fold abbreviated protocol showed a sensitivity of 98–100% and specificity of 99–100%, while the 4-fold protocol maintained a sensitivity of 95–98% and specificity of 99–100%. However, this paper focused on standard FSE sequences, and the whole protocol was repeated with 2 different compression levels; also, the DL protocols achieved a very high accuracy concerning the BME, but data regarding labrum abnormalities are limited in number [15]. The excellent diagnostic performance for labral lesion detection can be attributed to the isotropic voxel size (0.8 mm), enabling high-quality multiplanar reconstructions without step-ladder artifacts. This capability proves particularly valuable for the curved glenoid labrum geometry and obliquely oriented tears incompletely visualized on standard orthogonal planes [18–20]. The T1 shortening effect of diluted gadolinium creates a high signal within the joint space, providing excellent labral delineation that remains preserved despite accelerated acquisition and DL reconstruction. The superior performance of combined 3D protocols (Group A: 96.8% sensitivity) compared to single sequences (Group C: 87.5%, $p = 0.012$) demonstrates that T1 and PD-FS weightings provide complementary information. T1-weighted sequences optimize arthrographic effect and labral morphology, while PD-FS enhances bone marrow edema detection and extra-articular tissue contrast. This 7.3% accuracy improvement justifies acquiring both sequences despite the modest 3.5 min time penalty. The observed limitations for rotator cuff evaluation (75% sensitivity) align with theoretical expectations. Reduced contrast resolution between pathological and normal tendon substance, combined with the inability of arthrographic contrast to outline bursal-sided tears, likely explains the lower sensitivity. The paradoxically superior performance of PD-FS alone (Group C: 83.3%) for tendon evaluation confirms that fluid-sensitive contrast remains paramount for detecting subtle tendinopathy and partial tears. The near-real-time DL reconstruction (<1 min) represents a critical workflow advantage. The proprietary uAI algorithm's multiscale convolutional neural network, specifically trained on shoulder imaging data, preserves relevant anatomical details despite

2.5× acceleration factors. This rapid reconstruction ensures immediate image availability, maintaining clinical efficiency while achieving 75% total time reduction compared to standard protocols. The excellent inter-observer agreement ($\kappa > 0.8$) between readers of different experience levels (15 versus 4 years) suggests that established diagnostic criteria translate well to DL-accelerated 3D sequences. Our hypothesis of using 3D sequences either to complement or replace 2D FSE sequences is also supported by recent trends in the literature. A growing number of studies—particularly those focused on the knee—highlight the use of 3D imaging, with some of them exploring advanced technologies such as compressed sensing to significantly reduce acquisition times [22–24].

A critical consideration for clinical implementation concerns the implications of the slightly reduced sensitivity (95.1%) observed with DL-accelerated 3D sequences compared to conventional protocols. While this represents excellent overall performance, the 5–6% rate of missed labral tears translates to approximately 4–5 undetected lesions per 82–94 cases in our cohort. In the context of MR arthrography, where direct intra-articular gadolinium injection has been performed, repeating the examination using a conventional protocol if a tear is suspected but not detected is clinically impractical. Therefore, we advocate for a nuanced approach to clinical implementation. In high-risk instability patients or those with strong clinical suspicion for labral pathology, conventional multiplanar 2D protocols may remain preferable to minimize diagnostic uncertainty. Conversely, DL-accelerated 3D imaging may be most appropriate for: (1) screening examinations where clinical suspicion is lower, (2) patients with claustrophobia or difficulty tolerating prolonged scan times, (3) follow-up assessments where baseline lesions have been established, and (4) high-volume practices where the 75% time reduction significantly improves patient access. A hybrid approach—combining rapid 3D screening with selective targeted 2D sequences based on 3D findings—warrants prospective evaluation as a potential optimization strategy.

Study limitations include the absence of arthroscopic correlation, preventing the determination of absolute diagnostic accuracy. The reported sensitivity and specificity values therefore reflect agreement with standard 2D MRA interpretation rather than true diagnostic performance against surgical findings. The single-center design using one vendor's platform may limit generalizability, as different DL algorithms might yield varying results. Deep learning reconstruction algorithms are vendor-specific and trained on proprietary datasets, meaning that performance characteristics may differ substantially across platforms. Multi-center validation studies across different vendors are essential before these results can be generalized to broader clinical practice. Institutions considering the implementation of DL-accelerated protocols should ideally conduct site-specific validation to confirm diagnostic equivalence with their particular hardware and software configurations. The lack of comparison with non-accelerated 3D sequences prevents isolation of the DL acceleration effect from inherent 2D versus 3D differences. As a result, it is difficult to disentangle the diagnostic contribution of isotropic 3D imaging itself from that of the deep learning reconstruction.

Furthermore, our study design focused on the detection (presence/absence) of pathology rather than characterization of lesion extent or grading. Although the extent of bone marrow edema may not be critical for clinical decision-making, the extent of labral and rotator cuff tears is highly relevant for determining surgical candidacy. Future studies should incorporate structured grading systems for lesion extent to determine whether DL-accelerated 3D sequences can provide equivalent characterization to conventional protocols for pre-operative planning.

The difference in lesion identification between readers (94 vs. 82 labral lesions, 14.6% discrepancy) reflects expected inter-reader variability for subtle labral pathology, consistent with literature reports (κ typically 0.60–0.85 for labral assessment). Our per-

reader analytical design, where each reader's 3D performance was compared to their own 2D reference, provides a clinically meaningful assessment of diagnostic consistency when transitioning to accelerated protocols. However, definitive determination of whether the 12 discordant lesions represent true pathology would require arthroscopic correlation. This inter-reader variability underscores the inherent subjectivity in labral assessment and the importance of reader experience, particularly for borderline cases. The relatively lower sensitivity for bone marrow edema (82.9%) likely reflects compromised T2 weighting from shorter TR/TE values in accelerated sequences. This limitation should be considered when ordering examinations where bone marrow edema detection is clinically important, such as occult fracture evaluation or evaluation of rotator cuff abnormalities with associated findings.

The 75% time reduction improves patient comfort, reduces motion artifacts, and increases scanner throughput, potentially improving access to advanced imaging.

Future research should address arthroscopic correlation, multi-center validation across different platforms, and investigation of higher acceleration factors. Application to non-arthrographic shoulder MR and integration with automated diagnostic algorithms warrant exploration. The high-quality isotropic datasets could serve as ideal inputs for machine learning systems designed to detect shoulder pathology.

5. Conclusions

Deep learning-accelerated 3D MR arthrography sequences demonstrate high diagnostic accuracy for glenoid labral pathology while reducing acquisition time by 75%. Combined 3D protocols incorporating both T1 and PD-FS weightings optimize diagnostic performance, though conventional 2D sequences maintain advantages for subtle rotator cuff pathology and bone marrow edema. These findings support the selective implementation of DL-accelerated 3D protocols in shoulder MR arthrography, with conventional protocols remaining preferable for high-risk instability patients or those with strong clinical suspicion for labral pathology, where minimizing false-negative findings is paramount.

Author Contributions: Conceptualization, G.F. and P.A.; methodology, G.F. and P.A.; software, G.T., F.S., G.O. and L.M.; validation, G.T., G.F., P.A., F.S., G.O. and L.M.; formal analysis, F.S.; investigation, G.T. and G.O.; resources, G.T. and F.S.; data curation, G.F.; writing—original draft preparation, G.T. and F.S.; writing—review and editing, G.T., F.S. and G.F.; visualization, G.O. and L.M.; supervision, G.F. and P.A.; project administration, G.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of Comitato Etico delle Province di Verona e Rovigo; Project identification code Prot. n. 33264; Prog. 4287CESC; date of approval 31 May 2023.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The datasets generated and analyzed during the current study include individual patient imaging data and clinical information subject to patient privacy and ethical regulations. Due to the sensitive nature of medical imaging data and in accordance with institutional review board requirements, the raw imaging datasets are not publicly available to protect patient confidentiality. However, anonymized data supporting the findings of this study are available from the corresponding author upon reasonable request, subject to appropriate ethical approval and data sharing agreements. Requests will be reviewed by the institutional review board and must comply with applicable data protection regulations and patient consent provisions.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ACS	Accelerated Compressed Sensing
ALPSA	Anterior Labroligamentous Periosteal Sleeve Avulsion
BME	Bone Marrow Edema
BM	Bone Marrow
CAIPIRINHA	Controlled Aliasing in Parallel Imaging Results in Higher Acceleration
POLPSA	Posterior Labroligamentous Periosteal Sleeve Avulsion
PROPELLER	Periodically Rotated Overlapping Parallel Lines with Enhanced Reconstruction
SLAP	Superior Labrum Anterior to Posterior

References

1. Sebro, R.; Oliveira, A.; Palmer, W.E. MR arthrography of the shoulder: Technical update and clinical applications. *Semin. Musculoskelet. Radiol.* **2014**, *18*, 352–364. [\[CrossRef\]](#)
2. Llopis, E.; Montesinos, P.; Guedez, M.T.; Aguilera, L.; Cerezal, L. Normal Shoulder MRI and MR Arthrography: Anatomy and Technique. *Semin. Musculoskelet. Radiol.* **2015**, *19*, 212–230. [\[CrossRef\]](#) [\[PubMed\]](#)
3. De Maeseneer, M.; Van Roy, F.; Lenchik, L.; Shahabpour, M.; Jacobson, J.; Ryu, K.N.; Handelberg, F.; Osteaux, M. CT and MR arthrography of the normal and pathologic anterosuperior labrum and labral-bicipital complex. *Radiographics* **2000**, *20*, S67–S81. [\[CrossRef\]](#)
4. Magee, T.; Williams, D.; Mani, N. Shoulder MR arthrography: Which patient group benefits most? *AJR Am. J. Roentgenol.* **2004**, *183*, 969–974. [\[CrossRef\]](#) [\[PubMed\]](#)
5. De Filippo, M.; Schirò, S.; Sarohia, D.; Barile, A.; Saba, L.; Cella, S.; Castagna, A. Imaging of shoulder instability. *Skelet. Radiol.* **2020**, *49*, 1505–1523. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Daniels, S.P.; Greditzer, H.G., 4th; Mintz, D.N.; Dines, J.S.; Bogner, E.A. Swinging injuries in competitive baseball players. *Skelet. Radiol.* **2023**, *52*, 1277–1292. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Jung, J.Y.; Jee, W.H.; Park, M.Y.; Lee, S.Y.; Kim, Y.S. SLAP tears: Diagnosis using 3-T shoulder MR arthrography with the 3D isotropic turbo spin-echo space sequence versus conventional 2D sequences. *Eur. Radiol.* **2013**, *23*, 487–495. [\[CrossRef\]](#)
8. Beltran, J.; Bencardino, J.; Mellado, J.; Rosenberg, Z.S.; Irish, R.D. MR arthrography of the shoulder: Variants and pitfalls. *Radiographics* **1997**, *17*, 1403–1412; discussion 1412–1415. [\[CrossRef\]](#)
9. Xie, Y.; Tao, H.; Li, X.; Hu, Y.; Liu, C.; Zhou, B.; Cai, J.; Nickel, D.; Fu, C.; Xiong, B.; et al. Prospective Comparison of Standard and Deep Learning-reconstructed Turbo Spin-Echo MRI of the Shoulder. *Radiology* **2024**, *310*, e231405. [\[CrossRef\]](#)
10. Yang, W.; Zhang, X.; Tian, Y.; Wang, W.; Xue, J.H.; Liao, Q. Deep Learning for Single Image Super-Resolution: A Brief Review. *IEEE Trans. Multimed.* **2019**, *21*, 3106–3121. [\[CrossRef\]](#)
11. Hahn, S.; Yi, J.; Lee, H.J.; Lee, Y.; Lee, J.; Wang, X.; Fung, M. Comparison of deep learning-based reconstruction of PROPELLER Shoulder MRI with conventional reconstruction. *Skelet. Radiol.* **2023**, *52*, 1545–1555. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Cui, L.; Song, Y.; Wang, Y.; Wang, R.; Wu, D.; Xie, H.; Li, J.; Yang, G. Motion Artifact Reduction for Magnetic Resonance Imaging with Deep Learning and K-Space Analysis. *PLoS ONE* **2023**, *8*, e0278668. [\[CrossRef\]](#)
13. Yao, M.S.; Hansen, M.S. A Path Towards Clinical Adaptation of Accelerated MRI. *Proc. Mach. Learn. Res.* **2022**, *193*, 489–511. [\[PubMed\]](#)
14. 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*, 443–451. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Foti, G.; Spoto, F.; Mignolli, T.; Spezia, A.; Romano, L.; Manenti, G.; Cardobi, N.; Avanzi, P. Deep Learning-Driven Abbreviated Shoulder MRI Protocols: Diagnostic Accuracy in Clinical Practice. *Tomography* **2025**, *11*, 48. [\[CrossRef\]](#)
16. Johnson, P.M.; Lin, D.J.; Zbontar, J.; Zitnick, C.L.; Sriram, A.; Muckley, M.; Babb, J.S.; Kline, M.; Ciavarrà, G.; Alaia, E.; et al. Deep Learning Reconstruction Enables Prospectively Accelerated Clinical Knee MRI. *Radiology* **2023**, *307*, e220425. [\[CrossRef\]](#)
17. Reschke, P.; Gotta, J.; Gruenewald, L.D.; Bachir, A.A.; Strecker, R.; Nickel, D.; Booz, C.; Martin, S.S.; Scholtz, J.E.; D’Angelo, T.; et al. Deep Learning in Knee MRI: A Prospective Study to Enhance Efficiency, Diagnostic Confidence and Sustainability. *Acad. Radiol.* **2025**, *32*, 3585–3596. [\[CrossRef\]](#)
18. Soldatos, T.; Shah, J.P.; Chhabra, A. 3-Dimensional (3D) Isotropic MRI of the Shoulder—Advantages Over 2D MRI. *Semin. Roentgenol.* **2024**, *59*, 418–428. [\[CrossRef\]](#)
19. Foti, G.; Avanzi, P.; Mantovani, W.; Dal Corso, F.; Demozzi, E.; Zorzi, C.; Carbognin, G. MR arthrography of the shoulder: Evaluation of isotropic 3D intermediate-weighted FSE and hybrid GRE T1-weighted sequences. *La Radiol. Medica* **2017**, *122*, 353–360. [\[CrossRef\]](#) [\[PubMed\]](#)

20. Lee, S.H.; Yun, S.J.; Yoon, Y. Diagnostic performance of shoulder magnetic resonance arthrography for labral tears having surgery as reference: Comparison of high-resolution isotropic 3D sequence (THRIVE) with standard protocol. *La Radiol. Medica* **2018**, *123*, 620–630. [[CrossRef](#)]
21. Chang, P.D.; Chow, D.S. Revolutionizing Shoulder MRI: Accelerated Imaging with Deep Learning Reconstruction. *Radiology* **2024**, *310*, e233301. [[CrossRef](#)]
22. Wen, D.; Zhou, X.; Hou, B.; Zhang, Q.; Raithel, E.; Wang, Y.; Wu, G.; Li, X. 3D-DESS MRI with CAIPIRINHA two- and fourfold acceleration for quantitatively assessing knee cartilage morphology. *Skelet. Radiol.* **2024**, *53*, 1481–1494. [[CrossRef](#)] [[PubMed](#)]
23. Kijowski, R. 3D MRI of Articular Cartilage. *Semin. Musculoskelet. Radiol.* **2021**, *25*, 397–408. [[CrossRef](#)] [[PubMed](#)]
24. Oei, E.H.G.; van Zadelhoff, T.A.; Eijgenraam, S.M.; Klein, S.; Hirvasniemi, J.; van der Heijden, R.A. 3D MRI in Osteoarthritis. *Semin. Musculoskelet. Radiol.* **2021**, *25*, 468–479. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.