International Journal of Multiphysics
Volume 18, No. 3, 2024
ISSN: 1750-9548

SPARCC MRI Scoring in Axial Spondyloarthritis: A Comprehensive Review

Alaa Abd Elsamad Monge, Khaled Ahmed Lakouz, Essam Mohamed Hemat, Mohamed Abdelkhalek Basha, Fatma Ibrahim Abdelrahman Hassan

Department of Radiology, Faculty of Medicine, Zagazig University, Egypt

Corresponding author: Alaa Abd Elsamad Monge

Abstract

Background: Axial spondyloarthritis (axSpA) is a chronic inflammatory rheumatic disease primarily affecting the sacroiliac joints (SIJs) and spine, with significant diagnostic challenges due to nonspecific early symptoms and limited radiographic evidence. The Spondyloarthritis Research Consortium of Canada (SPARCC) MRI scoring system provides a standardized, reproducible method for quantifying bone marrow edema (BME) severity in SIJs.

Objective: This review synthesizes current evidence on the diagnostic utility of SPARCC MRI scoring, with particular emphasis on recent validation studies in understudied populations, optimal cutoff values for identifying high disease activity, and the unique epidemiological and immunogenetic considerations of the Middle East and North Africa (MENA) region.

Methods: This narrative review draws upon a recent prospective cross-sectional study conducted at Zagazig University, Egypt, alongside contemporary international literature on SPARCC validation, treat-to-target strategies, and regional disease characteristics.

Results: SPARCC scores demonstrate strong correlations with objective inflammatory markers (ESR, CRP) and composite disease activity indices (ASDAS-ESR, ASDAS-CRP), with the highest reported correlation being r = 0.929 for ASDAS-ESR. ROC analysis consistently yields excellent diagnostic accuracy (AUC > 0.94) for identifying high disease activity, with an optimal cutoff of ≥5 achieving 100% sensitivity, 77.3% specificity, and 91.7% accuracy. The MENA region exhibits markedly lower HLA-B27 positivity rates (11.7–65%) compared to Northern European populations (>90%), necessitating region-specific diagnostic algorithms.

Conclusion: SPARCC MRI scoring is a valid and clinically useful tool for assessing SIJ inflammation across diverse populations. The SPARCC ≥5 cutoff supports its integration into disease assessment and treatment decision-making, particularly in resource-limited settings where MRI access may be constrained. Future longitudinal studies are needed to validate treatment response monitoring and structural progression prediction in MENA populations.

Keywords: *axial spondyloarthritis, SPARCC, sacroiliac joint, MRI, bone marrow edema, HLA-B27, Middle East and North Africa, disease activity*

1. Introduction

Axial spondyloarthritis (axSpA) represents a spectrum of chronic inflammatory conditions primarily affecting the axial skeleton, encompassing both non-radiographic axSpA (nr-axSpA) and radiographic axSpA (ankylosing spondylitis, AS). The disease typically manifests in young adults with a male predominance and causes significant disability through progressive spinal inflammation and structural damage.

The diagnostic journey for axSpA patients is frequently prolonged, with international data documenting delays of 5–10 years between symptom onset and definitive diagnosis. This delay is particularly pronounced in

the MENA region, where limited access to specialized rheumatology care, restricted availability of MRI, and lower HLA-B27 prevalence contribute to diagnostic challenges.

Magnetic resonance imaging has emerged as the gold standard modality for detecting active inflammatory lesions in axSpA, capable of identifying bone marrow edema (BME)—the hallmark of active sacroiliitis—months to years before radiographic changes become apparent. Among the available MRI scoring systems, the SPARCC index has demonstrated particular utility due to its high reproducibility, sensitivity to change in longitudinal studies, and strong correlation with clinical disease activity measures.

2. The SPARCC MRI Scoring System: Methodological Foundations

2.1 Scoring Methodology

The SPARCC SIJ scoring system evaluates BME across six consecutive coronal STIR (short-time inversion recovery) sequences oriented parallel to the long axis of the sacrum. The methodology follows ten standardized steps:

1. Dichotomous scoring: All lesions are scored as present (1) or absent (0).
2. Slice selection: Six coronal slices (typically 4–9) representing the largest proportion of the synovial compartment are assessed.
3. Sequence specificity: Only STIR sequence abnormalities are scored; T1-weighted images serve anatomical reference.
4. Anatomical boundaries: Iliac bone lesions are scored in their entirety; sacral lesions are scored medially to the lateral border of the sacral foramina.
5. Reference standard: Sacral inter-foraminal bone marrow signal establishes the threshold for normal versus increased signal.
6. Quadrant division: Each SIJ is divided into four quadrants (upper/lower iliac; upper/lower sacral), yielding a maximum of 8 points per slice.
7. Intensity scoring: "Intense" signal (referenced against presacral venous blood) adds 1 point per joint per slice (max 12).
8. Depth scoring: "Deep" lesions extending ≥1 cm from the articular surface add 1 point per joint per slice (max 12).
9. Blinded assessment: Pre- and post-treatment images are scored together with the observer blinded to time sequence.
10. Calibration: Reference cases and control images are available for proficiency training.

The total maximum SPARCC score is 72: 48 points for BME presence, 12 for intense edema, and 12 for deep edema.

2.2 Validation and Reliability

The SPARCC system has undergone extensive validation since its introduction by Maksymowych and colleagues. Recent work has focused on enhancing scoring proficiency through the SPARCC MRI-RETIC e-tools, which provide calibration modules for both inflammatory and structural lesions. These modules require scoring of 20–50 MRI cases with achievement of intraclass correlation coefficient (ICC) targets for status and change scores. In the EuroSpA validation study, calibration using RETIC modules significantly improved reliability between readers and SPARCC developer radiologists, with ICC values improving substantially after training, particularly for readers with limited prior SPARCC experience.

System usability assessments indicate that both the inflammatory (RETIC-INF) and structural (RETIC-STR) modules score above the widely-accepted threshold of 68 on the System Usability Scale, suggesting they are suitable for widespread clinical and research application.

3. Correlations with Clinical and Laboratory Parameters

3.1 Inflammatory Markers

The relationship between SPARCC scores and systemic inflammatory markers has been consistently demonstrated across multiple cohorts, though with some heterogeneity explained by baseline inflammatory burden. In the Egyptian validation study, SPARCC scores exhibited exceptionally strong correlations with both ESR ($r = 0.870$, $p < 0.0001$) and CRP ($r = 0.857$, $p < 0.0001$). These values exceed those reported in several international cohorts, likely attributable to the higher baseline inflammatory burden in this population (median CRP 29 mg/L; median ESR 73.5 mm/h).

This finding aligns with meta-analytic evidence encompassing 55 studies that validated the MRI–inflammatory marker relationship across diverse cohorts. However, discordant results from cohorts with lower baseline inflammatory markers (e.g., median CRP ~17.5 mg/L) highlight an important methodological consideration: range restriction in inflammatory markers can obscure true associations. Cohorts with broader inflammatory variability are more likely to detect significant correlations.

3.2 Disease Activity Indices

A hierarchical pattern of correlations has emerged consistently across studies, as summarized in Table 1.

Table 1. Hierarchical pattern of SPARCC correlations with disease activity indices

Correlation Pair	Typical Range	Interpretation
SPARCC–BASDAI	r = 0.53–0.65	Moderate
SPARCC–VAS pain	r = 0.66–0.70	Moderate-to-strong
SPARCC–ASDAS	r = 0.74–0.80	Strong
SPARCC–ASDAS-CRP	r = 0.80–0.86	Strong
SPARCC–ASDAS-ESR	r = 0.88–0.93	Very strong

The moderate SPARCC–BASDAI association reflects the multifactorial nature of BASDAI, which incorporates elements beyond active inflammation (structural damage, central sensitization, psychological comorbidities). The exceptionally strong SPARCC–ASDAS-ESR correlation ($r = 0.929$ in the Egyptian cohort) is among the highest reported and has practical implications: in resource-limited settings, ASDAS-ESR may serve as a reliable, low-cost surrogate for MRI-detected inflammation.

3.3 Functional Measures

Both BASFI (Bath Ankylosing Spondylitis Functional Index) and BASMI (Bath Ankylosing Spondylitis Metrology Index) demonstrate strong positive correlations with SPARCC scores ($r \approx 0.77–0.78$), indicating that greater SIJ inflammation is associated with worse functional impairment and reduced spinal mobility. This relationship appears robust even in established disease (>4 years duration), suggesting that ongoing inflammatory burden continues to contribute to functional limitation alongside structural damage—a finding that supports aggressive inflammation suppression through biologic DMARDs.

4. Diagnostic Performance and Optimal Cutoff Values

4.1 ROC Analysis

ROC curve analysis for identifying high disease activity (ASDAS ≥ 2.1) has consistently demonstrated excellent diagnostic accuracy, as summarized in Table 2.

Table 2. ROC analysis for identifying high disease activity across studies

Study Population	AUC	95% CI
Egyptian cohort (Zagazig University)	0.943	0.874–1.000
International validation (Maksymowych et al.)	0.960	—
Egyptian simplified SPARCC validation	0.794	—

The Egyptian comprehensive study's AUC of 0.943 closely aligns with international benchmarks and substantially exceeds the simplified validation study, likely reflecting differences in reference standard and cohort disease severity.

4.2 Cutoff Value Analysis

Systematic evaluation of multiple SPARCC thresholds reveals a clear trade-off between sensitivity and specificity, as presented in Table 3.

Table 3. Diagnostic validity of SPARCC MRI score at various cutoff values. *Optimal threshold (maximized Youden's index).

Cutoff	Sensitivity	Specificity	PPV	NPV	Accuracy
≥2	100%	27.3%	70.4%	100%	73.3%
≥3	100%	50.0%	77.6%	100%	81.7%
≥4	100%	72.7%	86.4%	100%	90.0%
≥5*	100%	77.3%	88.4%	100%	91.7%

The SPARCC ≥5 threshold emerges as optimal, maximizing Youden's index while maintaining perfect sensitivity. This cutoff is concordant with recommendations by Weber et al. and Maksymowych et al., who advocated for higher SPARCC thresholds when confirming high inflammatory activity to justify treatment escalation. Notably, the ASAS classification threshold of SPARCC ≥2 (indicating definite active sacroiliitis) achieves perfect sensitivity but poor specificity (27.3%), rendering it suitable for classification but inadequate for confirming high disease activity requiring therapeutic intervention.

5. Regional Considerations: The MENA Context

5.1 HLA-B27 Epidemiology

A critical finding from the Egyptian study—and one with profound diagnostic implications—is the markedly low HLA-B27 positivity rate of 11.7%, which stands in sharp contrast to the 90–95% prevalence reported in Northern European populations.

Regional data reveal substantial heterogeneity, as summarized in Table 4.

Table 4. HLA-B27 prevalence across regions

Region/Country	HLA-B27 Positivity in axSpA	General Population Prevalence
Northern Europe	90–95%	6–25%
Egypt (Zagazig cohort)	11.7%	~2–5%
Turkey	15–40%	—
Iran	10–30%	—

MENA pooled	25–75%	0.3–7%
Qatar (Arab ethnicity)	74–82%	—

This variability reflects the complex interplay of genetic subtypes, consanguinity patterns, and environmental factors. The ASAS-PerSpA study demonstrated that while a positive family history of axSpA was associated with HLA-B27 positivity across most global regions, this association was notably absent in the MENA region—further underscoring the distinct immunogenetic architecture of axSpA in this population.

5.2 Diagnostic Algorithm Implications

The low HLA-B27 positivity in Egyptian and broader MENA populations carries critical practical consequences:

1. Over-reliance on HLA-B27 testing would systematically exclude the majority of true axSpA patients from appropriate diagnostic pathways.
2. MRI assumes heightened importance as a diagnostic modality, particularly in nr-axSpA where radiographic sacroiliitis is absent.
3. Quantitative scoring systems (SPARCC) provide objective benchmarks that can standardize diagnostic and treatment decisions across heterogeneous populations.
4. Access challenges remain significant, as MRI availability is limited in many MENA healthcare systems, and the 40–60 minute scoring time per patient constrains scalability.

5.3 Disease Phenotype in MENA

The Egyptian cohort exhibited several distinctive features:

• Long disease duration (mean 27.1 years), reflecting substantial diagnostic delays

• Uniformly high disease activity: median BASDAI 7, mean ASDAS-CRP 3.7, mean ASDAS-ESR 4.21

• Elevated acute-phase reactants: median CRP 29 mg/L, median ESR 73.5 mm/h

• High SPARCC scores: median 11 (range 1–14), indicating very high SIJ inflammatory activity

These characteristics suggest that by the time of diagnosis, MENA patients often present with advanced, highly active disease—reinforcing the need for earlier detection strategies and streamlined referral pathways.

6. SPARCC in the Era of Treat-to-Target

The treat-to-target (T2T) paradigm has gained traction in axSpA management, with clinical remission (ASDAS < 1.3) serving as the primary target. The TRACE study—a recent multicenter trial evaluating secukinumab-based T2T strategies—provides important longitudinal context for SPARCC utility.

In this multicenter trial, SPARCC and CANDEN scoring were used to assess SIJ and spine inflammation at baseline and weeks 4, 16, 24, and 56. Key findings included:

• Rapid inflammation reduction: Total MRI inflammation (SIJ + spine) decreased from mean 39.0 to 6.7 over 56 weeks

• MRI remission definition: SPARCC-SIJ ≤ 1 + CANDEN-spine ≤ 2 achieved in 43% at week 56

• Structural remodelling: Erosion decreased while fat lesions and backfill increased by week 4, with new bone formation/ankylosis remaining unchanged

• Predictors of MRI remission: Lower baseline MRI inflammation and HLA-B27 positivity

Notably, changes in MRI inflammation and structural lesions were comparable irrespective of clinical remission

status at week 56, suggesting that MRI and clinical remission are not perfectly concordant—an important consideration when using SPARCC as a treatment response biomarker.

7. Practical Challenges and Future Directions

7.1 Current Limitations

Despite its proven value, SPARCC scoring faces practical barriers:

• Time intensity: 40–60 minutes per patient limits feasibility in high-volume setting

• Training requirements: Specialized radiologist expertise is needed for reliable scoring

• Inter-reader variability: Though improved by RETIC calibration, formal reliability assessments remain essential

• Cross-sectional constraint: Current validation is largely cross-sectional; longitudinal treatment response data in MENA populations are sparse

7.2 Emerging Innovations

Several developments promise to address these limitations:

1. Deep learning algorithms: Automated SPARCC scoring using convolutional neural networks is under active development, with preliminary studies demonstrating feasibility for SIJ MRI assessment.

2. Simplified scoring variants: The Egyptian simplified SPARCC validation (AUC 0.794) suggests that abbreviated scoring methods may offer reasonable accuracy with reduced time burden, though at some cost to diagnostic precision.

3. RETIC e-tool expansion: Continued refinement of online calibration modules is improving reader reliability globally, with particular benefit for regions with limited local SPARCC expertise.

4. Integration with artificial intelligence: Machine learning approaches may eventually enable real-time SPARCC estimation during MRI acquisition, potentially streamlining clinical workflows.

7.3 Research Priorities

Priority areas for future investigation include:

• Longitudinal validation of the SPARCC ≥5 cutoff for treatment monitoring in MENA cohorts
• Structural progression prediction: Whether baseline SPARCC scores predict future radiographic damage
• Cost-effectiveness analyses of MRI-SPARC-guided versus clinically-guided treatment algorithms in resource-limited settings

• Inter-reader reliability studies specifically in MENA radiology practices

• Pediatric and juvenile-onset axSpA SPARCC validation

8. Conclusion

The SPARCC MRI scoring system represents a robust, validated tool for quantifying sacroiliac joint inflammation in axial spondyloarthritis. Its strong correlations with objective disease activity measures, excellent diagnostic performance (AUC 0.943), and identification of an optimal cutoff value (SPARCC ≥5) for high disease activity support its integration into comprehensive disease assessment protocols.

For the MENA region specifically, this evidence is particularly salient. The markedly lower HLA-B27 positivity rates, prolonged diagnostic delays, and high baseline disease activity documented in Egyptian cohorts underscore the urgent need for objective, imaging-based diagnostic benchmarks that do not rely on genetic markers. SPARCC scoring fulfills this need while providing a quantitative framework for treatment decision-making and response monitoring.

As treat-to-target strategies become increasingly central to axSpA management, and as automated scoring technologies mature, SPARCC is poised to transition from a specialized research tool to a routine clinical instrument—provided that training infrastructure and quality assurance mechanisms are established, particularly in underserved regions. The evidence synthesized here provides a foundation for such implementation, while

highlighting the critical importance of population-specific validation in ensuring equitable, effective care for axSpA patients worldwide.

References

1. Maksymowych WP, Hadsbjerg AEFEF, Østergaard M, et al. Validation of SPARCC MRI-RETIC e-tools for increasing scoring proficiency. RMD Open. 2024;10:e003923.

2. Lin Y, Cao P, Chan SCW, et al. Deep Learning Algorithm of the SPARCC Scoring System in SI Joint MRI. J Magn Reson Imaging. 2024. DOI: 10.1002/jmri.29211.

3. Vladimirova N, et al. Clinical and MRI remission in axial spondyloarthritis: one-year outcomes and predictors from the TRACE treat-to-target study with secukinumab. Arthritis Res Ther. 2026;18:38.

4. El Zorkany B, et al. The treatment journey for patients with axial spondyloarthritis in North Africa and the Middle East: From diagnosis to management. Rheumatol Ther. 2020;7:829-844.

5. Khudadah M, et al. Managing spondyloarthritis in Middle East and North Africa. Arab J Rheumatol. 2024.

6. Reveille JD. How Disease Severity, Ethnicity, and HLA-B27 Prevalence Intersect. Spondylitis Plus. 2024.

7. Boel A, et al. Geographical prevalence of family history in patients with axial spondyloarthritis. Ann Rheum Dis. 2022;80(Suppl 1):341.

8. Weber U, Østergaard M, Lambert RG, et al. Candidate lesion-based criteria for defining a positive sacroiliac joint MRI in axSpA. Ann Rheum Dis. 2015;74:1976-1982.

9. Maksymowych WP, et al. Defining the Minimally Important Change for the SPARCC Spine and Sacroiliac Joint MRI Indices. J Rheumatol. 2012;39:1666-1674.