Journal of Radiation Research and Imaging

Lal B, et al., J Radiat Res Imaging. 2025;3(1):9-13.

Commentary

Advancing AI in MRI interpretation for temporomandibular disorders: a new era of diagnosis

Babu Lal, MDS, FDSRCPS, MAMS[1,*], Hariram Sankar, MDS[2], Ragavi Alagarsamy, MDS[3], Ankush Ankush, MD[4]

[1]Assistant Professor, Department of Trauma and Emergency Medicine, All India Institute of Medical Sciences, Bhopal, Madhya Pradesh, India

[2]Assistant Professor, Department of Dental Surgery, PSG Institute of Medical Sciences and Research, Coimbatore, India

[3]Project Research Scientist-I, Department of oral and Maxillofacial surgery, Maulana Azad institute of dental sciences, New Delhi, India

[4]Assistant Professor, Department of Radiodiagnosis and Imaging, LNMC & JK Hospital, Bhopal, India

*Author for correspondence:
Email: meethariram205@gmail.com

Received date: December 08, 2025
Accepted date: December 30, 2025

Copyright: © 2025 Lal B, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Introduction

Temporomandibular disorders (TMDs) have long sat at the uncomfortable intersection between diagnostic ambiguity and therapeutic uncertainty [1–3]. Magnetic Resonance Imaging (MRI) has been the gold standard for assessing articular disc position, morphology, and early soft-tissue changes [4–6]. Yet despite its value, interpretation remains dependent on expert skill, institutional resources, and considerable time investment.

Sankar *et al.*'s recent systematic review on the role of artificial intelligence (AI) in MRI-based detection of TMD disc position comes at a pivotal time [7]. Radiology is in the midst of a paradigm shift where automated tools are no longer aspirational—they are being validated, deployed, and in some domains, trusted for first-pass triage. In oral and maxillofacial radiology, however, AI adoption has lagged compared to fields like neuroradiology or thoracic imaging [8–12]. This commentary provides one of the first structured syntheses of how deep learning models can identify disc displacement on MRI with high sensitivity and specificity. But the conversation cannot stop at detection. The real question is how AI can move from static image interpretation to clinical decision support, personalized care, and workflow transformation. This commentary explores that broader landscape and highlights where the next steps in progress are likely to come from.

Understanding the Diagnostic Landscape of TMD

Clinical heterogeneity and diagnostic variability

TMDs are multifactorial [13]. Disc displacement, while frequently observed, does not always correlate with clinical symptoms. Several longitudinal studies have shown that many cases of anterior disc displacement with reduction remain stable or asymptomatic for years. This weak structure–symptom relationship complicates diagnosis and contributes to the chronic under- or over-treatment of TMD. MRI provides a visual anchor to this uncertainty. It helps distinguish normal disc-condyle relationships from anterior displacement, disc perforation, or effusion. But radiologic diagnosis is not equivalent to clinical diagnosis. Even among experienced radiologists, inter-observer variability remains a concern. Subtle changes in mouth opening angle, MRI sequence parameters, or image contrast can yield different interpretations.

Where AI fits in this picture

AI's greatest strength is consistency. Trained on large datasets, deep learning models can identify patterns across thousands of scans with the same decision rules every time. The systematic review highlighted that models like MobileNet V2 and ResNet achieved accuracy above 83% and, in some studies, sensitivities approaching 1.0 [14]. This level of reproducibility is difficult to match with

human readers, particularly in lower-resource settings. However, these models currently operate in isolation from the clinical context. They detect presence or absence of displacement—but not significance, symptom correlation, or treatment implications. Bridging that gap is where the next phase of research must focus.

The Technological Backbone: Promise and Pitfalls

Deep learning vs. classical ML

Of the seven studies included in the review, six used deep learning models and one used classical machine learning (Random Forest) [14–20]. The latter performed poorly compared to Convolutional Neural Network (CNN)-based architectures. This reflects a broader trend: classical ML requires manual feature engineering and performs inconsistently on complex imaging tasks, whereas CNNs (e.g., ResNet, U-Net, MobileNet V2) learn discriminative features directly from the data. The performance metrics of all the seven included studies were illustrated in **Figure 1**. But not all CNNs are created equal. Architectural choice influences not just performance metrics but also interpretability, data requirements, and deployment feasibility. For example, U-Net excels at segmentation tasks but may require clear ROI delineation [15–19]. ResNet handles large datasets well and is robust to overfitting. MobileNet V2 is lightweight and suitable for real-time clinical applications, especially in resource-limited settings.

Data quality and heterogeneity

The accuracy of any model depends on the quality of its training data. In the reviewed studies, MRI protocols varied widely: different mouth positions, different sequences (proton density, T1, T2), and inconsistent labelling standards. This heterogeneity limits model generalizability. A network trained on closed-mouth proton density sequences from one center may fail miserably on T2-weighted open-mouth sequences from another. The solution is not more data alone—it's standardized data. Federated learning, as mentioned by the authors, offers one path forward. By training models on distributed datasets without sharing patient information, institutions can contribute to large, diverse training cohorts while respecting privacy.

ROI definition and preprocessing

Region of interest (ROI) extraction was a recurring challenge. The TMJ is a small structure, and disc visualization occupies a tiny fraction of the MRI field. Automatic or semi-automatic ROI detection improves model efficiency, reduces noise, and allows networks to focus on clinically relevant anatomy. Data augmentation (rotations, zooming, CLAHE contrast enhancement) helped compensate for small datasets in several studies. This is an important lesson: careful preprocessing can sometimes rival dataset expansion in its impact on performance.

Figure 1. Performance metrics and characteristics of the included studies.

Clinical Integration

The problem with detection-only models

Imagine a model that flags anterior disc displacement on every MRI it sees. It would be accurate, fast, and utterly unhelpful without context. Clinicians don't treat images; they treat patients. Whether a displaced disc warrants intervention depends on symptoms, duration, functional limitation, and sometimes psychosocial factors [13–21]. For AI to have clinical impact, it must integrate, Clinical metadata (e.g., pain scores, range of motion, history), Additional imaging markers (e.g., effusion, condylar flattening), and temporal evolution (e.g., follow-up scans).

Decision support versus automation

The realistic role of AI in TMJ imaging is decision support, not replacement. Think of it as a tireless junior resident: pre-screens the images, highlights regions of interest, and provides structured outputs. The final judgment rests with the clinician. This model preserves accountability and reduces diagnostic delays.

Addressing disparities in access

One of the most powerful implications of lightweight architectures like MobileNet V2 is democratization of expertise. In settings where TMJ radiologists are scarce, AI-assisted interpretation could standardize diagnostic quality, enable earlier diagnosis, reduce misdiagnosis in peripheral hospitals, and guide referral pathways more intelligently. This is not speculative; similar models have already transformed diabetic retinopathy screening and tuberculosis detection in low-resource settings.

Methodological Gaps and Reporting Standards

Lack of prospective and multicentric validation

All seven studies in the review were retrospective. None involved prospective clinical evaluation, let alone randomized trials. Retrospective models often perform well on internal datasets but collapse when exposed to real-world variability. External validation—especially across geographic, hardware, and population variations, is essential before clinical deployment. Nozawa *et al*. made an important contribution here by demonstrating domain shift: models trained at one site underperformed when tested on external data. That finding should be a wake-up call for anyone assuming lab performance equals clinical readiness.

Lack of standardized outcome reporting

Sensitivity, specificity, accuracy, and AUC were variably reported across studies. Some used open versus closed mouth comparisons, others didn't. Some stratified by diagnosis, others aggregated all TMDs. Without consistent metrics, it's difficult to benchmark models or meta-analyze results. An international reporting standard tailored to AI in maxillofacial imaging could accelerate progress. Such a framework might require, Clear description of MRI protocols, dataset size and split, ROI definition methods, model architecture, validation strategy (internal vs external), and full confusion matrix reporting.

Absence of cost-effectiveness and workflow studies

No study evaluated economic implications. Yet in healthcare systems, cost matters as much as accuracy. AI that saves radiologist time, shortens report turnaround, or enables non-expert interpretation can have outsized economic impact. Without cost modelling, adoption remains theoretical.

Ethical and Regulatory Considerations

Data privacy and federated learning

TMD MRI datasets are inherently small at single centers. Pooling data across institutions is the only way to train robust models, but patient privacy regulations make this challenging. Federated learning offers a way to train on distributed data without moving it. This could enable large-scale training across academic and community hospitals, particularly in regions like India and East Asia where TMD prevalence is high.

Algorithmic bias

Most of the included studies came from East Asian populations. Craniofacial morphology varies across populations. Models trained on these datasets may not generalize to other ethnic groups. Without attention to algorithmic bias, AI risks amplifying disparities rather than reducing them.

Clinical Scenarios and Use Cases

Triage in high-volume imaging centers

AI could pre-classify scans into "likely normal" versus "possible displacement," allowing radiologists to prioritize abnormal cases. This mirrors successful chest X-ray triage models used in TB screening.

Decision support for early-career clinicians

Surgeons and radiologists in training often struggle with disc visualization, particularly in low-contrast images. An AI-generated heat map highlighting the disc-condyle interface could serve as a real-time training tool.

Telemedicine and peripheral deployment

MobileNet-type architectures can be embedded in low-resource MRI workstations. This could allow primary or district hospitals to send AI-flagged cases for remote expert review, improving equity of care.

Beyond MRI: Expanding the Role of AI Across the TMD Diagnostic Ecosystem

Here's the thing: if we keep talking about AI in TMD only through the lens of MRI disc detection, we're missing most of the story. The joint doesn't fail in one dimension. Bone, soft tissue, function, symptoms, and even biomechanics all play a part. What's happening in the recent literature shows that AI is starting to tie these threads together in ways we simply couldn't before.

CBCT-based AI: bringing the hard-tissue side into focus

MRI gives us the disc, but it tells us very little about the bony architecture that often drives symptoms or long-term degeneration. Cone-Beam computed tomography (CBCT) fills that gap, and AI is beginning to make CBCT interpretation far more powerful.

Mourad *et al*. showed that a YOLO-based model can pick up TMJ-Osteo arthritis (OA) features—erosion, cysts, flattening—with Area Under the curve (AUCs) hovering around 0.87–0.91 and agreement

levels that nearly match expert examiners [22]. Eser *et al.* report something similar: YOLOv5 hit perfect sensitivity for segmentation and performed reliably across multiple (OA) phenotypes on CBCT sagittal slices. What this really means is that AI is starting to give us automated, consistent, bone-level assessment—the side of TMD we usually diagnose too late [23].

Predicting soft-tissue problems from CBCT alone

Choi *et al.* extend the conversation by testing whether CBCT radiomics can offer meaningful clues about disc displacement. Their Random Forest models reached an AUC of 0.86 for predicting DDWOR, the best performance in their set of experiments. It's not enough to replace MRI, but it does give clinicians in MRI-limited settings a more informed way to triage and decide who should be referred for advanced imaging [23].

Clinical-only AI models: when symptoms tell the story

MRI and CBCT are only part of the puzzle. Clinical features—pain intensity, maximum mouth opening, lateral excursion, psychological scores—carry a ton of diagnostic signal. Yıldız *et al.* showed that a model trained only on these practical clinic-floor measurements reached an AUC of 0.9387. It's a quiet reminder that some of the smartest AI tools don't need an image at all. They just need structured, high-quality clinical data [24].

Functional AI: movement, occlusion, and real-world joint behavior

Farook *et al.* highlight a rapidly growing area: AI for jaw tracking, occlusal analysis, and motion prediction using chairside sensors, video, 3D scanners, and virtual articulators. Static images freeze the joint in time. Functional AI shows how the system behaves. For TMD, that's game-changing [25].

Paediatric TMJ: AI filling a real diagnostic gap

Paediatric TMJ imaging is notoriously difficult—small discs, limited cooperation, and motion everywhere. Azma *et al.*'s two-step U-Net++ model hits solid segmentation performance and identifies pediatric disc displacement with AUCs up to 0.92. For clinicians uncomfortable reading pediatric TMJ MRI, AI can level the playing field [26].

AI that makes MRI itself better

Not all AI in imaging is about diagnosis. Some of them make the images themselves more usable.

AI-assisted compressed sensing (ACS): Ye *et al.* cut scan times nearly in half—13:28 down to 6:08—while improving Signal-to-Noise Ratio and Contrast-to-Noise Ratio and still hitting strong diagnostic agreement for disc displacement and effusion [27,28].

Deep learning reconstruction (DLR): Jo *et al.* found Deep Learning (DL)-reconstructed MRI to be diagnostically interchangeable with conventional MRI, with almost 50% shorter scan duration and noticeably cleaner images [29]. Shorter scans mean fewer repeats, fewer motion artifacts, and more patient comfort—especially in painful open-mouth positions.

From pattern recognition to mechanism: AI meets biomechanics

The most intriguing shift may come from work like Sun *et al.*, who combine explainable DL with multiscale biomechanical modelling [30]. Their system identifies morphological risk factors for TMD—small mandibular size, flat condylar shape—and then simulates how these factors alter joint forces, nutrient diffusion, and disc strain. This moves AI past "detecting abnormalities" and toward explaining why the joint fails. It's the kind of insight that could eventually drive targeted prevention, not just earlier diagnosis.

A Roadmap for Future Research

Standardization and open science

- Establish standard MRI protocols for TMD imaging (including mouth position, sequence parameters, and reference planes).
- Develop public benchmark datasets with annotated disc positions, enabling fair comparison across models.
- Encourage code and model sharing under appropriate privacy frameworks.

Clinical trials and validation

- Design prospective multicentric trials evaluating model performance in real-world workflows.
- Include clinical endpoints such as change in diagnostic turnaround time, inter-observer variability reduction, and treatment planning impact.
- Evaluate cost-effectiveness and patient outcomes.

Ethical AI development

- Ensure diversity in training datasets.
- Incorporate explainability modules.
- Adhere to international reporting and regulatory standards.

Conclusion: From Algorithms to Action

Sankar *et al.*'s systematic review does more than aggregate accuracy scores—it marks an inflection point in how we might approach TMD diagnostics [7]. The message is clear: AI can match, and sometimes exceed, human performance in identifying disc displacement on MRI. But detection alone is not enough. The future lies in clinically integrated, explainable, and validated AI systems that augment clinician expertise rather than replace it. This means moving from image classification to context-aware decision support, from retrospective studies to prospective trials, and from academic silos to collaborative federated networks.

AI's role in TMD imaging could mirror what it's already achieving in fields like oncology and cardiology: transforming complex, subjective interpretation into structured, rapid, and equitable care delivery. Realizing that vision will require technical rigor, clinical insight, and regulatory clarity—but the groundwork is being laid right now.

Acknowledgements

None.

Conflict of Interest

The author declares no conflicts of interest.

References

1. Dym H, Israel H. Diagnosis and treatment of temporomandibular disorders. Dent Clin North Am. 2012 Jan 1;56(1):149.
2. Gauer RL, Semidey MJ. Diagnosis and treatment of temporomandibular disorders. Am Fam Physician. 2015 Mar 15;91(6):378–86.
3. Reid KI, Greene CS. Diagnosis and treatment of temporomandibular disorders: an ethical analysis of current practices. J Oral Rehabil. 2013 Jul;40(7):546–61.
4. Kaur H, Singh R, Sankar H, Vichare S. Magnetic Resonance Imaging for Craniofacial Applications. InApplications of Three-dimensional Imaging for Craniofacial Region 2024 Oct 19 pp. 35–52.
5. Larheim TA. Role of magnetic resonance imaging in the clinical diagnosis of the temporomandibular joint. Cells Tissues Organs. 2005;180(1):6–21.
6. Vogl TJ, Lauer HC, Lehnert T, Naguib NN, Ottl P, Filmann N, et al. The value of MRI in patients with temporomandibular joint dysfunction: Correlation of MRI and clinical findings. Eur J Radiol. 2016 Apr;85(4):714–9.
7. Sankar H, Alagarsamy R, Lal B, Rana SS, Roychoudhury A, Barathi A, et al. In reply to the Letter to Editor regarding Comment on: Role of artificial intelligence in magnetic resonance imaging-based detection of temporomandibular joint disorder: a systematic review. Br J Oral Maxillofac Surg. 2025 Jun;63(5):407–9.
8. Kalyanpur A, Mathur N. Applications of artificial intelligence in thoracic imaging: a review. Acad Med. 2025 Feb 21;2(1).
9. Chassagnon G, Vakalopoulou M, Paragios N, Revel MP. Artificial intelligence applications for thoracic imaging. Eur J Radiol. 2020 Feb;123:108774.
10. Hwang EJ, Goo JM, Park CM. AI Applications for Thoracic Imaging: Considerations for Best Practice. Radiology. 2025 Feb;314(2):e240650.
11. Gong B, Khalvati F, Ertl-Wagner BB, Patlas MN. Artificial intelligence in emergency neuroradiology: Current applications and perspectives. Diagn Interv Imaging. 2025 Apr;106(4):135–42.
12. Carvalho Macruz FB, Dias ALMP, Andrade CS, Nucci MP, Rimkus CM, et al. The new era of artificial intelligence in neuroradiology: current research and promising tools. Arq Neuropsiquiatr. 2024 Jun;82(6):1–12.
13. Sankar H, Lal B, Rana SS, Mishra RK. The common Clinical Characteristics and Etiology of Temporomandibular Disorder in a Tertiary Care Center: An Observational Study. J Maxillofac Oral Surg. 2025 Jul 28:1–7.
14. Yoon K, Kim JY, Kim SJ, Huh JK, Kim JW, Choi J. Explainable deep learning-based clinical decision support engine for MRI-based automated diagnosis of temporomandibular joint anterior disk displacement. Comput Methods Programs Biomed. 2023 May;233:107465.
15. Kao ZK, Chiu NT, Wu HH, Chang WC, Wang DH, Kung YY, T, et al. Classifying Temporomandibular Disorder with Artificial Intelligent Architecture Using Magnetic Resonance Imaging. Ann Biomed Eng. 2023 Mar;51(3):517–26.
16. Ozsari S, Yapicioglu FR, Yilmaz D, Kamburoglu K, Guzel MS, Bostanci GE, et al. Interpretation of magnetic resonance images of temporomandibular joint disorders by using deep learning. IEEE Access. 2023 May 18;11:49102–13.
17. Lee YH, Won JH, Kim S, Auh QS, Noh YK. Advantages of deep learning with convolutional neural network in detecting disc displacement of the temporomandibular joint in magnetic resonance imaging. Sci Rep. 2022 Jul 5;12(1):11352.
18. Lin B, Cheng M, Wang S, Li F, Zhou Q. Automatic detection of anteriorly displaced temporomandibular joint discs on magnetic resonance images using a deep learning algorithm. Dentomaxillofac Radiol. 2022 Mar 1;51(3):20210341.
19. Nozawa M, Ito H, Ariji Y, Fukuda M, Igarashi C, Nishiyama M, et al. Automatic segmentation of the temporomandibular joint disc on magnetic resonance images using a deep learning technique. Dentomaxillofac Radiol. 2022 Jan 1;51(1):20210185.
20. Orhan K, Driesen L, Shujaat S, Jacobs R, Chai X. Development and Validation of a Magnetic Resonance Imaging-Based Machine Learning Model for TMJ Pathologies. Biomed Res Int. 2021 Jul 5;2021:6656773.
21. Schiffman E, Ohrbach R, Truelove E, Look J, Anderson G, Goulet JP, et al. Diagnostic Criteria for Temporomandibular Disorders (DC/TMD) for Clinical and Research Applications: recommendations of the International RDC/TMD Consortium Network* and Orofacial Pain Special Interest Group†. J Oral Facial Pain Headache. 2014 Winter;28(1):6–27.
22. Mourad L, Aboelsaad N, Talaat WM, Fahmy NMH, Abdelrahman HH, El-Mahallawy Y. Automatic detection of temporomandibular joint osteoarthritis radiographic features using deep learning artificial intelligence. A Diagnostic accuracy study. J Stomatol Oral Maxillofac Surg. 2025 Sep;126(4):102124.
23. Eşer G, Duman ŞB, Bayrakdar İŞ, Çelik Ö. Classification of temporomandibular joint osteoarthritis on cone beam computed tomography images using artificial intelligence system. J Oral Rehabil. 2023 Sep;50(9):758–66.
24. Choi E, Kim D, Lee JY, Park HK. Artificial intelligence in detecting temporomandibular joint osteoarthritis on orthopantomogram. Sci Rep. 2021 May 13;11(1):10246.
25. Yıldız NT, Kocaman H, Yıldırım H, Canlı M. An investigation of machine learning algorithms for prediction of temporomandibular disorders by using clinical parameters. Medicine (Baltimore). 2024 Oct 11;103(41):e39912.
26. Farook TH, Dudley J. Understanding Occlusion and Temporomandibular Joint Function Using Deep Learning and Predictive Modeling. Clin Exp Dent Res. 2024 Dec;10(6):e70028.
27. Azma R, Hareendranathan A, Li M, Nguyen P, S Wahd A, Jaremko JL, et al. Automated pediatric TMJ articular disk identification and displacement classification in MRI with machine learning. J Dent. 2025 Apr;155:105622.
28. Ye Z, Lyu X, Zhao R, Fan P, Yang S, Xia C, et al. Accelerated MRI in temporomandibular joints using AI-assisted compressed sensing technique: a feasibility study. Eur Radiol. 2025 Dec;35(12):8046–57.
29. Ye Z, Xiong X, Lv MX, Li Z. AI-assisted Compressed Sensing MRI in Temporomandibular Joints: A Feasibility Study. J Med Imaging Radiat Sci. 2024 Oct 1;55(3):101610.
30. Jo GD, Jeon KJ, Choi YJ, Lee C, Han SS. Deep learning reconstruction for temporomandibular joint MRI: diagnostic interchangeability, image quality, and scan time reduction. Eur Radiol. 2025 Sep 25.
31. Sun S, Xu P, Buchweitz N, Hill CN, Ahmadi F, Wilson MB, Mei A, She X, Sagl B, Slate EH, Lee JS, Wu Y, Yao H. Explainable deep learning and biomechanical modeling for TMJ disorder morphological risk factors. JCI Insight. 2024 Jul 11;9(16):e178578.