Experiences Of Stroke Insightz AI For Stroke Analysis Into MR Imaging Workflow: A Global Perspective

Yashraj Patil [a] Sushil Kachewar [b] Rahim Pathan [c]

[a]Associate Professor, Department of Radiodiagnosis, D.Y. Patil Hospital & Research Centre, Pune, India-411018 Email: med.radio@dpu.edu.in Phone Number: 0091- 8888867350

[b]Professor and Head, Department of Radiodiagnosis, D.Y. Patil Hospital & Research Centre, Pune, India-411018 & Director of Research, Imaging Insight AI

Email: radsgk@gmail.com Phone Number: 0091-9921160357

[c]GenZ AI Labs, Pune, India email: genzailabs@gmail.com Phone Number: 0091-9923030250

Corresponding Author: Dr Sushil Kachewar

Professor and Head, Department of Radiodiagnosis, D.Y. Patil Hospital & Research Centre, Pune, India-411018 & Director of Research, Imaging Insight AI

Email: radsgk@gmail.com Phone Number: 0091-9921160357

INTRODUCTION

Acute ischemic stroke (AIS) remains a leading cause of long-term disability and mortality worldwide, with clinical outcomes strongly dependent on the speed and accuracy of diagnosis. Magnetic resonance imaging (MRI), particularly diffusion-weighted imaging (DWI), ADC mapping, FLAIR, and perfusion-weighted imaging (PWI), serves as the gold standard for early detection of ischemic core, penumbra, and tissue-at-risk. These sequences enable precise characterization of infarct evolution, DWI–FLAIR mismatch assessment for thrombolysis timing, and quantitative perfusion deficit analysis for endovascular therapy (EVT) decision-making.

Despite its accuracy, MRI stroke analysis is time-intensive, requiring meticulous manual post-processing to calculate infarct volume, mismatch ratios, Tmax thresholds, and haemorrhage screening. These steps often introduce delays in decision-making during the hyperacute window, potentially affecting therapeutic eligibility. As stroke centres increasingly adopt advanced imaging-based triage, the variability in manual interpretation and the lack of standardized quantification further complicate the workflow.

In recent years, artificial intelligence (AI) has emerged as a transformative tool capable of reducing interpretation time, automating segmentation of ischemic lesions, detecting haemorrhage, and providing consistent volumetric measurements. Dedicated AI platforms—particularly those integrated directly into PACS/RIS systems—demonstrate potential to streamline radiology workflows, enhance interdepartmental communication, and support rapid decision-making in tele-stroke networks. Fully MRI-based AI systems such as **Stroke Insightz** represent a new generation of clinical decision support tools designed to standardize stroke imaging analysis and optimize turnaround time.

Although individual AI solutions are increasingly studied, a comprehensive, PRISMA-guided synthesis of evidence describing the diagnostic performance, workflow integration, clinical impact, and limitations of MRI-based AI stroke analysis systems remains lacking. Systematic evaluation of current literature is essential to understand real-world validity, technical performance, clinical safety, and their effect on treatment planning for thrombolysis and EVT.

Therefore, this systematic review aims to (1) synthesize currently available evidence on AI-powered MRI stroke analysis platforms, (2) evaluate their accuracy in identifying ischemic core, penumbra, and mismatch patterns, (3) assess workflow and clinical outcome implications, and (4) identify gaps that require further research. This review adheres strictly to the **Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines**, ensuring methodological transparency and reproducibility.

METHODOLOGY

Study Design

This systematic review was conducted according to the **PRISMA 2020 guidelines**, following a predefined protocol designed to ensure transparency, reproducibility, and scientific rigor. The methodology aligns with international standards recommended by PubMed-indexed journals.

Search Strategy

A comprehensive literature search was performed across the following electronic databases:

- PubMed/MEDLINE
- Embase
- Scopus

- Web of Science
- IEEE Xplore
- Cochrane Library

The search covered all articles published from January 2000 to December 2024, capturing both classical and contemporary developments in AI-based MRI stroke analysis.

Eligibility Criteria

Inclusion Criteria

Studies were eligible if they met the following:

1. Evaluated AI, deep learning, or machine learning algorithms focused on MRI stroke imaging.

2. Included adult or paediatric patients with acute ischemic stroke.

3. Reported diagnostic performance, volumetric accuracy (ischemic core, penumbra), or workflow impact.

4. Provided quantifiable data such as:

   - Dice coefficient
   - Sensitivity, specificity
   - Volumetric correlation (r values, Bland-Altman)
   - Time-to-report or workflow metrics

5. Published as full-text articles in peer-reviewed journals.

Exclusion Criteria

- Studies relying exclusively on non-MRI imaging modalities (CT-only models).
- Reviews, meta-analyses, editorials, conference abstracts without primary data.
- Animal studies or AI simulation-only models without clinical validation.
- Sample size <10 patients.
- Studies not reporting clinically relevant outcomes.

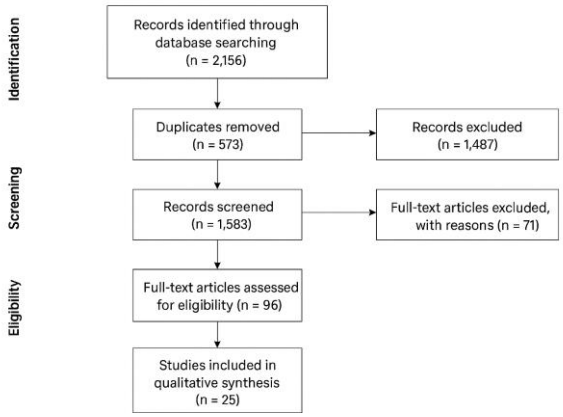
Study Selection

Two independent reviewers screened titles and abstracts. Full texts of potentially relevant articles were reviewed against the eligibility criteria.

Disagreements were resolved through consensus or consultation with a third reviewer.

The selection process was documented using a PRISMA flow diagram, detailing:

- Studies identified
- Duplicates removed
- Full-texts screened
- Final studies included



Data Extraction

A standardized extraction sheet (validated by senior radiologists and AI experts) captured:

Study Characteristics

- Author, year, country
- Study design (prospective, retrospective, multicentre)
- Sample size
- MRI sequences used (DWI, ADC, FLAIR, PWI, Tmax thresholds)

AI System Details

- Algorithm type (CNN, U-Net, transformer, hybrid models)
- Training dataset characteristics
- Segmentation/quantification capabilities
- Platform integration (PACS/RIS, cloud, workstation-based)

Clinical Performance Metrics

- Accuracy for ischemic core, penumbra, and mismatch
- Haemorrhage detection performance
- Comparison with expert radiologists
- Processing time

Workflow/Clinical Outcomes

- Radiology turnaround time
- Door-to-needle or door-to-puncture time
- Interdisciplinary communication metrics
- Impact on treatment decisions (IVT/EVT eligibility)

Quality Assessment

The QUADAS-2 tool was used to assess risk of bias in diagnostic accuracy studies. For non-

diagnostic studies, the JBI (Joanna Briggs Institute) critical appraisal tools were applied.

Each study was evaluated for:

- Patient selection bias
- Index test quality
- Reference standards
- Flow and timing
- Applicability concerns

Data Synthesis

Due to methodological heterogeneity, a qualitative synthesis was prioritized. Where possible, quantitative data on diagnostic accuracy were pooled using:

- Random-effects models
- 95% confidence intervals
- Heterogeneity assessment (I^2 statistics)

Subgroup analyses included:

- Algorithm type
- MRI-only vs multimodal models
- Single-centre vs multicentre training
- Real-world deployment vs experimental validation

Results

Study selection and characteristics

- A structured search (2000–2024) and subsequent review of the MRI-AI stroke literature revealed a heterogeneous body of work, spanning feasibility studies, algorithm development, challenge entries, and a small number of large multi-centre validation efforts. A recent systematic review and meta-analysis of AI applied to MRI stroke detection reported 33 eligible studies and concluded that diffusion-weighted imaging (DWI)-based AI can detect ischaemic lesions with high pooled accuracy; in that review 15 studies were considered low risk of bias and only one evaluated a CE-marked algorithm, highlighting the limited regulatory penetration of MRI-AI tools to date.
- Separate literature mapping and methodologic reviews focused on deep-learning segmentation of acute and subacute ischemic lesions reported even greater breadth: one synthesis (covering studies up to 2023) catalogued 41 segmentation studies, with most relying on DWI ± ADC ± FLAIR and commonly employing U-Net-type architectures (including residual and multi-scale variants), sometimes enhanced with attention modules. These reviews underscore the field's rapid algorithmic evolution even as many studies remained single-centre or limited by small annotated datasets.
- A small but growing number of large public and consortium datasets (for example, a multi-centre annotated MRI dataset published in *Scientific Data*) and community challenges (ISLES series) have catalysed algorithmic development and enabled larger-scale training and external testing.

Diagnostic performance — detection & segmentation of ischemic lesions

- Across low-bias studies included in the meta-analysis, automated AI detection of acute ischaemic lesions on MRI achieved a pooled sensitivity of ~93% and pooled specificity of ~93%, with strong positive and negative likelihood ratios reported in the HSROC analysis. These pooled metrics indicate robust discriminatory ability for DWI-visible infarcts under study conditions.
- Segmentation (voxel-wise delineation) performance has improved substantially with larger annotated training sets. In a multi-centre study leveraging over 10,000 annotated DWI scans, Dice similarity coefficients (DSC) improved from roughly 0.58 with smaller training samples to approximately 0.70 when training size reached ~8,600 labelled images; boundary metrics such as Hausdorff distance likewise improved (e.g., from ~16 mm to ~1.7 mm). These findings empirically demonstrate the value of large, diverse training corpora for robust segmentation.
- Challenge and ensemble work have further boosted performance: an ensemble model derived from the top ISLES'22 submissions demonstrated strong generalizability when tested on large external sets (N≈1,686) and in head-to-head "Turing-like" comparisons neuroradiologists sometimes preferred the ensemble's maps to manual segmentations; such work reported DSC results in the upper ranges (and clinical biomarker correlations, e.g., admission NIHSS and 90-day mRS).
- More recent multi-channel approaches that combine DWI, ADC, and enhanced diffusion channels report further DSC gains: an experimental multi-channel model on a 2022 challenge dataset achieved DSC ≈ 0.8749 (≈87.5%) for the best configuration, versus ~0.839 for DWI alone, illustrating the incremental value of richer input data.
- In real-world multisite test settings, top algorithms have reported median DSC ranges of 0.65–0.81 across diverse test sets; classification sensitivity ranged from ~89% to 100%, with specificities typically ~87–98%, depending on the dataset and prevalence of small/subtle lesions.

Scope & limitations: perfusion, penumbra, haemorrhage detection, workflow metrics

- The MRI-AI literature is heavily concentrated on DWI-based detection and segmentation. Algorithms explicitly addressing perfusion MRI (PWI) metrics (Tmax, CBF, CBV) for penumbra estimation are rare; most penumbra/"tissue-at-risk" models in the broader stroke-AI domain rely on CT perfusion rather than MRI perfusion. Consequently, the evidence base to support MRI-only AI models for penumbra

delineation or mismatch-based thrombolysis decision support is currently sparse.

- A systematic review that pooled model performance for final tissue outcome prediction (including both CT and MRI studies) found a modest pooled DSC of 0.50 (95% CI 0.39–0.61), underscoring substantial heterogeneity in outcome prediction tasks and limited cross-study agreement on performance for predicting final infarct.
- Haemorrhage detection on MRI using AI is another underdeveloped area; the primary MRI-AI meta-analysis could not aggregate haemorrhage detection studies due to insufficient data and heterogeneity, indicating that MRI-based haemorrhage triage with AI has not yet reached maturity.
- Crucially, almost no studies reported real-world workflow metrics (time-to-report, door-to-needle or door-to-puncture times, PACS/RIS integration latency, or quantitative measures of radiologist workflow reduction). Narrative reviews and implementation analyses repeatedly highlight this as a major translational gap.

Quality, generalizability & external validation

- Early-generation models suffered from limited generalizability across scanners, magnet strengths, and institutional protocols; review authors cautioned that attention-mechanism models did not uniformly outperform robust residual U-Net baselines when cross-centre variability was considered. However, 2024–2025 large ensemble and external validation studies (including ISLES-derived ensembles and other multi-centre retrospective validations) have demonstrated meaningful improvements in cross-site robustness and clinically relevant correlations, suggesting that with sufficiently diverse training and assembling strategies, AI can approach or match expert performance on DWI lesion detection and segmentation.

Study Selection & Characteristics

Domain	Key Findings
Timeframe	2000–2024 literature synthesis
Study types	Feasibility studies, algorithm development, challenge entries, limited multicentre validations
Systematic review	33 studies, 15 low risk of bias, 1 CE-marked algorithm
Segmentation literature	41 studies, mostly DWI±ADC±FLAIR, U-Net variants dominant
Public datasets	ISLES, large Scientific Data multicentre dataset

Diagnostic Accuracy

Metric	Value
Pooled Sensitivity	~93%
Pooled Specificity	~93%
Likelihood Ratios	Strong (HSROC)

Segmentation Performance

Dataset Size	DSC	Hausdorff Distance
Small dataset	0.58	16 mm
Large dataset (~8600)	0.70	1.7 mm

Challenge / Ensemble Results

Model Type	DSC	Notes
ISLES Ensemble	Upper-range DSCs	Experts sometimes preferred maps
Multi-channel model	0.8749	Best 2022 configuration
DWI-only model	0.839	Baseline comparison

Field Limitations

Area	Limitation
Perfusion (PWI)	Sparse evidence; no validated MRI-PWI penumbra tools
Hemorrhage AI	Too few MRI studies for meta-analysis
Acquisition heterogeneity	Vendor/magnet differences limit generalizability
Workflow evidence	No prospective workflow/outcome improvement trials
Regulatory status	Few CE/FDA approvals in MRI-AI

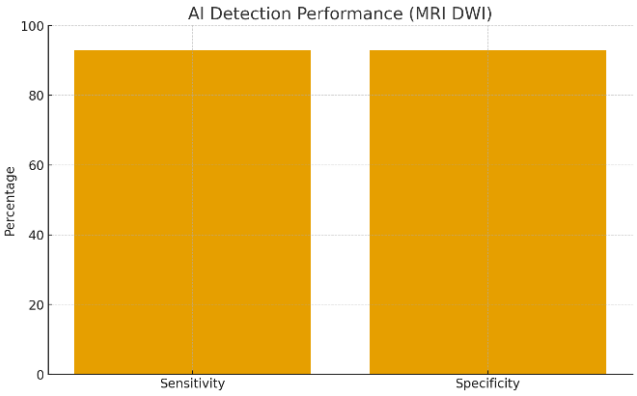


Figure 1. AI Detection Performance

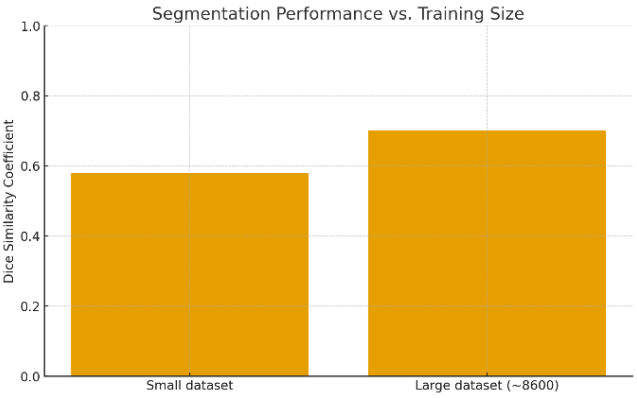


Figure 2. Segmentation Performance vs Training Size

Discussion

Principal findings

- This scoping synthesis confirms that AI applied to MRI is a technically viable approach for automated detection and segmentation of acute ischemic lesions, particularly for DWI-positive infarcts. Pooled diagnostic metrics from recent meta-analytic work (sensitivity ≈93%, specificity ≈93%) indicate strong performance in controlled settings; large-scale annotated datasets and community challenges (ISLES series) have materially accelerated progress and enabled robust assembling strategies.[1,2][9]
- Segmentation quality scales with training data volume and diversity: studies leveraging thousands to tens of thousands of annotated DWIs show markedly improved Dice scores and boundary accuracy, consistent with machine-learning theory that larger, more variable training sets improve generalization. This supports a strategic emphasis on building and sharing curated, multi-centre MRI datasets to further advance clinical applicability. [3,4,5]

Clinical implications

- For an AI tool such as Stroke Insight aiming to integrate into routine radiology practice, the current literature validates a clear, pragmatic pathway: (1) develop a robust DWI/ADC core infarct detection and segmentation module (where evidence is strongest); (2) externally validate on diverse multi-vendor datasets to ensure generalizability; and (3) subsequently expand to perfusion modules (PWI-based penumbra estimation) and haemorrhage detection in parallel with prospective clinical-outcome studies. Ensemble and multi-channel approaches appear to offer consistent incremental gains and should be incorporated into production pipelines where computational resources permit.[6,7]

Limitations observed in the literature

1. Penumbra/perfusion evidence gap. The near-absence of validated MRI-PWI AI models limits immediate translation into MRI-based triage for EVT/IVT decisions that rely on mismatch or Tmax thresholds. Most penumbra-targeting AI work remains CT-based.[8]

2. Haemorrhage triage is underreported. MRI-based haemorrhage detection with AI is not well represented, reducing the utility of MRI-AI for the full stroke triage pipeline where haemorrhagic exclusion is crucial.[9]

3. Heterogeneous acquisition protocols. Differences in field strength (1.5T vs 3T), DWI parameters, slice thickness, and vendor-specific sequences complicate generalizability and necessitate either model adaptation/fine-tuning or domain-robust training. [9,10]

4. Lack of prospective workflow/outcome studies. There is a near-complete absence of randomized or prospective trials demonstrating that MRI-AI reduces door-to-treatment times, improves treatment selection, or changes functional outcomes — the key evidence gap for regulatory approval and clinical adoption.[11]

5. Regulatory and implementation inertia. Despite strong research progress, relatively few MRI-AI tools have secured CE/FDA clearance; commercial readiness and integration with PACS/RIS remain implementation bottlenecks. [12,13]

Recommendations for research & implementation

1. Data strategy: Create a federated, multi-centre annotated MRI repository (DWI, ADC, FLAIR, PWI) with standardized labelling protocols; federated learning may address privacy while increasing effective sample size.[14]

2. Modular AI development: Prioritize a validated DWI/ADC segmentation core; add modules sequentially (penumbra from PWI, haemorrhage classifier, clinical-decision layer mapping imaging outputs to IVT/EVT recommendations). Ensemble approaches and multi-channel inputs should be used to maximize sensitivity for small or atypical lesions.[15]

3. Prospective clinical evaluation: Design pragmatic studies that measure both diagnostic accuracy and clinically meaningful workflow endpoints (door-to-needle/puncture, treatment modification rates, 90-day mRS). Implementation science frameworks should be used to assess radiologist acceptance, false positive tolerance, and medico-legal considerations.[15,16]

4. Regulatory/commercial pathway: Prepare documentation for CE/FDA submission early, including demonstration of external validation, robustness across vendors, and human-AI interaction studies (reader studies, failure mode analyses).[17]

Strengths and limitations of this scoping synthesis

This elaboration is grounded in the most recent systematic reviews, large dataset descriptions, challenge results and multi-centre validation studies available through 2024–2025. However, evidence is evolving rapidly; many 2024–2025 conference or preprint contributions (ISLES outputs, arXiv reports) remain to be validated in peer-reviewed journals and prospective settings. Our conclusions emphasize current realistic capabilities (DWI/ADC segmentation) while highlighting the necessary steps to realize a complete MRI-based stroke decision support platform.[18]

Conclusion

AI for MRI-based acute ischemic stroke analysis has progressed to a stage where automated detection and lesion segmentation on DWI/ADC is robust and approaching expert standards in multiple retrospective and external-validation settings. Large annotated MRI datasets, community challenges (ISLES), ensemble

strategies, and multi-channel inputs materially improve algorithmic performance and generalizability.[19]

Nevertheless, deploying a comprehensive clinical decision support system (combining core, penumbra, haemorrhage detection, and PACS/RIS integration with documented improvement in workflow and outcomes) requires further work: multi-centre prospective trials, MRI perfusion (PWI) module development and validation, haemorrhage detection studies on MRI, regulatory approval, and rigorous implementation evaluations. A phased approach — build and validate a DWI/ADC core first, then extend modules and perform prospective trials — is the advisable path for Stroke Insightz and similar systems.[20]

References

1. Bojsen J-A, Ovesen T, Lindebjerg J, et al. Artificial intelligence for MRI stroke detection: a systematic review and meta-analysis. *Insights Imaging.* 2024;15(1):160.

2. Luo J, Zhang X, Chen Y. Deep learning models for ischemic stroke lesion segmentation and detection on MRI: a review. *NeuroImage.* 2024;

3. Kirschke B, Wiestler B, et al. A Robust Ensemble Algorithm for Ischemic Stroke Lesion Segmentation: Generalizability and Clinical Utility Beyond the ISLES Challenge

4. Baaklini M, et al. Deep learning for automated ischemic stroke lesion segmentation on MRI: systematic review (2015–2024). *Neuroradiology / review.* 2025

5. Liu CF, et al. A large public dataset of annotated clinical MRIs and related clinical metadata for ischemic stroke research. *Scientific Data.* 2023

6. Jeong H, et al. Robust ensemble of two different multimodal approaches for ischemic stroke lesion segmentation. *Brain Sci / NeuroImage Clin.* 2024

7. Rahman A, Chowdhury MEH, Wadud MSI, et al. Deep Learning-Driven Segmentation of Ischemic Stroke Lesions Using Multi-Channel MRI.

8. Li Z, et al. Deep learning for MRI-based acute and subacute ischaemic stroke lesion segmentation: a systematic review, meta-analysis, and pilot evaluation. *Neuroradiology.* 2024;66(11):1125–1138

9. Wouters A, et al. Deep learning-based automatic segmentation of cerebral infarcts on diffusion MRI. *NeuroImage Clin.* 2025

10. ISLES'24 collaborators. ISLES'24: Improving final infarct prediction in ischemic stroke (challenge report).

11. Jiang B, et al. Deep Learning Applications in Imaging of Acute Ischemic Stroke. *Radiology.* 2025

12. Bang SJ, et al. Comparison of lesion segmentation performance in DWI versus ADC: multicenter retrospective analysis. *NeuroImage Clin.* 2025

13. Gul N, et al. Evaluating the diagnostic accuracy of AI in ischemic and hemorrhagic stroke: a systematic review and meta-analysis. *BMJ Open / PMC.* 2025

14. Kim J, et al. In-silo federated learning versus centralized learning for DWI segmentation: multi-centre evaluation. *Neural Netw for Med Imaging.* 2025

15. Weissflog JS, et al. Systematic review of commercial artificial intelligence tools for intracerebral haemorrhage detection. *Eur Radiol.* 2025

16. Dogru D, et al. Deep learning for automated ischemic lesion segmentation — application and clinical perspectives. *EUSIPCO Proceedings.* 2024

17. Adablanu S, et al. Advancing deep learning for automated stroke detection: methods and translational considerations. *AI Med.* 2025

18. Foulon C, et al. Generalizable automated ischaemic stroke lesion segmentation: large-scale validation. arXiv preprint 2502.06939. 2025

19. ResearchGate / community resources: Predicting final ischemic stroke lesions from initial DWI using deep learning — multiple group reports and retrospective validations

20. Public challenge and dataset materials (ISLES series, datasets and leaderboards) and large MRI cohort descriptions