# Initial Insights Into a Secure Large Language Model for MRI Examination Requests: A Retrospective Study

James Thomas Patrick Decourcy Hallinan[1,2], MBChB; Naomi Wenxin Leow[3], BComp, MComp; Yi Xian Low[1], MBBS; Aric Lee[1], MBBS; Wilson Ong[1], MBBS; Matthew Ding Zhou Chan[1], BmedSc, MD; Gordan Toh Cheong Zheng[5], Diploma in Biomedical Science; Ganakirthana Kalpenya Devi[1], MbBchBao; Stephanie Shengjie He[1], MBBS; Daniel De-Liang Loh[1], MBBS, MRCS; Desmond Shi Wei Lim[1], MBBS; Xi Zhen Low[1], MBBS; Mei Chin Lim[1], MBBS; Clement Yong[1], MBBS; Weizhong Jonathan Sng[1], MBBS; Ee Chin Teo[1], MMRT; Jiong Hao Tan[4], MBBS, MMed, MRCS; Naresh Kumar[4], MBBS, MS, DNB, DM; Andrew Makmur[1,3], BmedSc, MBBS, MMed; Yonghan Ting[1], MBBS

[1] Dept. of Diagnostic Imaging, NUH, Singapore
[2] Dept. of Diagnostic Radiology, YLL School of Medicine, NUS, Singapore
[3] AI Office, NUHS, Singapore
[4] National University Spine Institute, NUH, Singapore
[5] Republic Polytechnic, Singapore

Correspondence:
Dr James Hallinan, Senior Consultant, Department of Diagnostic Imaging, National University Hospital
Email: james_Hallinan@nuhs.edu.sg

## Background

- Incomplete Magnetic Resonance Imaging (MRI) requests lead to suboptimal protocols, repeat studies & reduced accuracy.
- LLMs can interpret clinical text at scale, but privacy constraints limit direct clinical use
- Institutional sLLMs operate behind firewalls, protecting patient data while enabling augmentation

## Objectives

For all MRI subspecialty modalities (Musculoskeletal, Neuroradiology, body):

1. Does sLLM augmentation improve clinical information quality of MRI examination requests?
2. Can sLLM match radiologist-level MRI protocol accuracy?

## Materials and Methods

**Figure 1:** Summary of materials

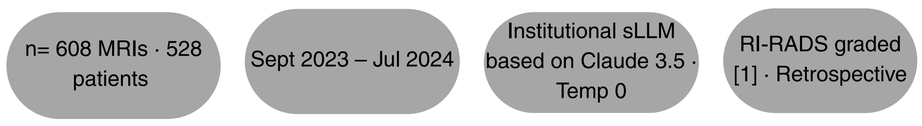
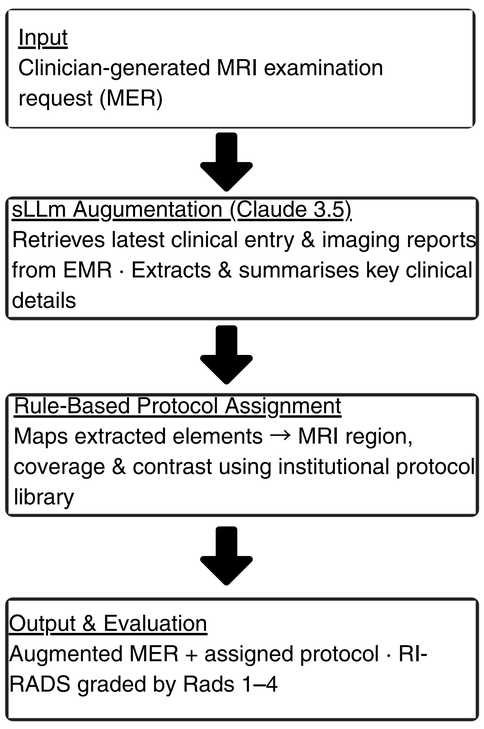
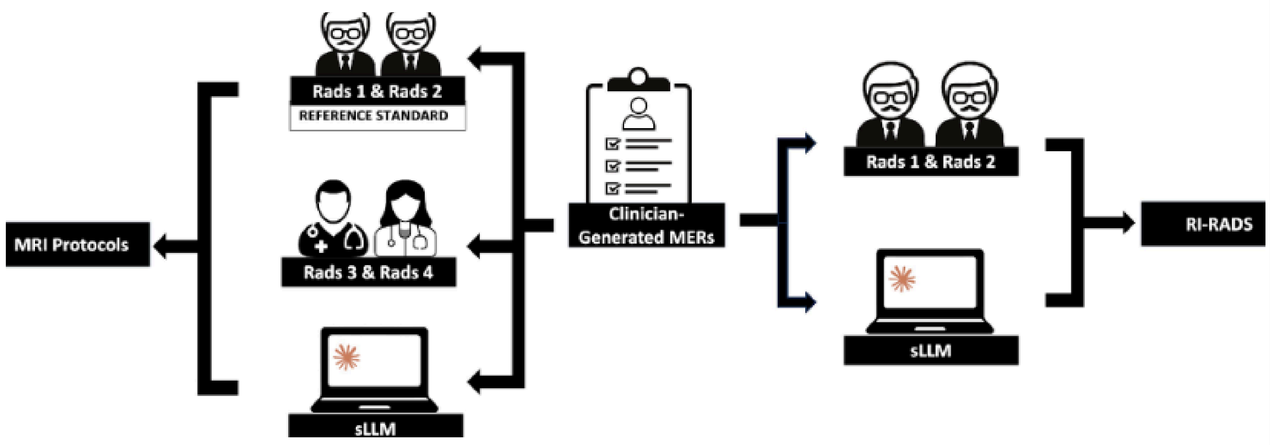



**Figure 2:** Flowchart of sLLM augmentation



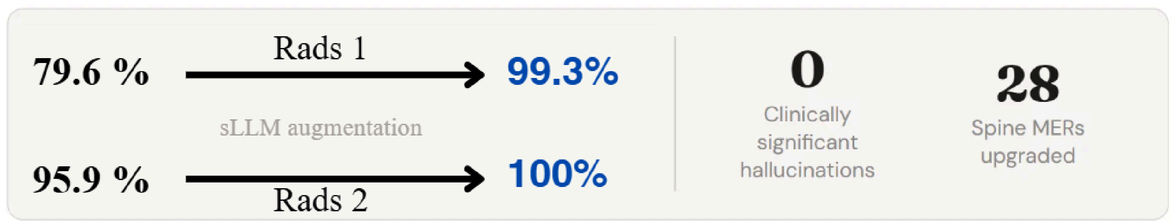

| Grade | Description | Information included in the requisition |
|---|---|---|
| **RI-RADS A** | **Adequate** | All key categories of information included. |
| **RI-RADS B** | **Barely adequate** | All key categories of information included, some clinical findings missing. |
| **RI-RADS C** | **Considerably limited** | Two categories of information included. |
| **RI-RADS D** | **Deficient** | One or no category of information included. |

**Figure 3:** Study design flowchart. Clinician MERs were augmented using an sLLM (Claude 3.5, Anthropic). Two experienced radiologists (Rads 1 and 2) graded information quality via RI-RADS and provided the protocol accuracy reference standard. The sLLM and two general radiologists (Rads 3 and 4) were evaluated against this standard.




## Main Findings

**Figure 4:** Summary of results of MRI protocoling accuracy for the sLLM and board-certified radiologists versus reference standard. sLLM augmentation nearly eliminated inadequate MRI requests and matched expert radiologist protocol accuracy.



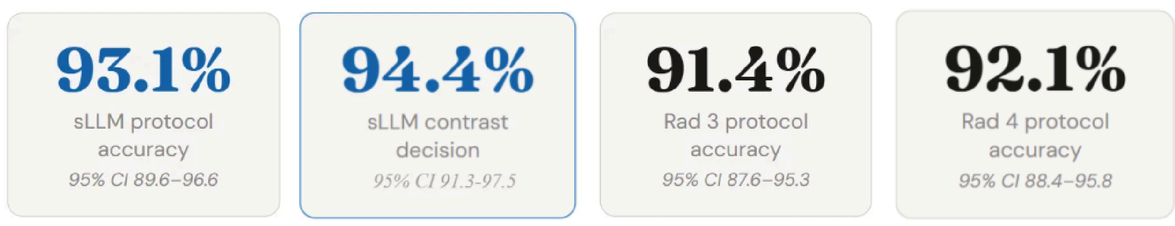




**Figure 5:** Overall MRI Protocoling accuracy for the sLLM and board-certified radiologists versus reference standard.

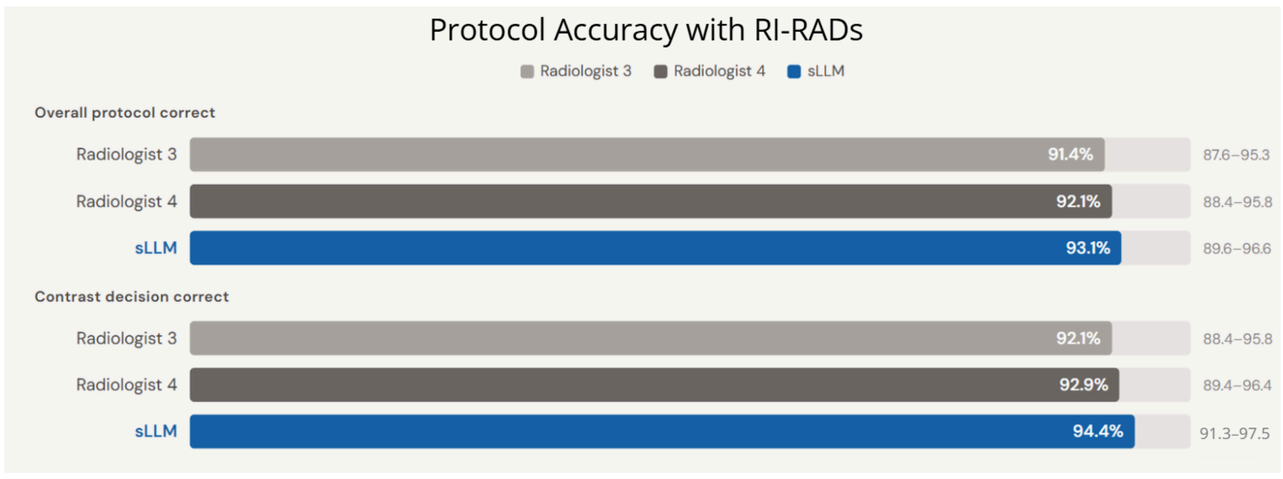
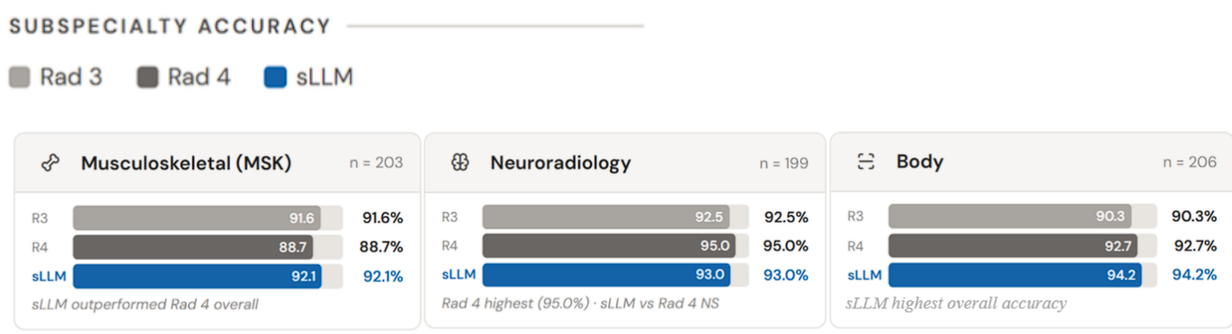



**Figure 6:** MRI protocoling accuracy for the sLLM and board-certified radiologists versus reference standard, broken down into subspecialties




## Discussion

- sLLM upgraded MER quality across all subspecialties, fewer than 1% remained inadequate post-augmentation
- sLLM protocol accuracy (93.1%) was comparable to board-certified radiologists (Rad 3: 91.4%, Rad 4: 92.1%); consistent with prior LLM studies in spine/MSK [2] and CT protocol assignment [3].
- Critical clinical details absent from original MERs (surgery history, tumour, infections) were extracted and summarised by the sLLM
- Rule-based script made final protocol decisions — sLLM provided clinical context only, not the decision

## Limitations

- Single-centre, retrospective study · outpatient MERs only
- Not powered for non-inferiority testing
- Post-deployment guardrails remain essential

## Conclusion

- Near 100% adequate MERs after sLLM augmentation vs 79.6-95.9% clinician-only
- Protocol accuracy matched experienced radiologists across body, MSK & neuroradiology
- End-to-end pipeline offers a scalable, standardised protocoling pathway with reduced admin burden
- Future work: prospective multicentre evaluation, cost-effectiveness & extension to CT / ultrasound

References

1. Wildman-Tobriner B, et al. RI-RADS: A New Standardized Reporting System for the Reason for Exam. J Am Coll Radiol. 2022.
2. Hitzert MM, et al. Large Language Models for the Automated Vetting of Musculoskeletal MRI Requests. J Digit Imaging. 2024.
3. Kanemaru N, et al. Efficacy of Fine-Tuned Large Language Model in CT Protocol Assignment as Clinical Decision-Supporting System. J Imaging Inform Med. 2025.