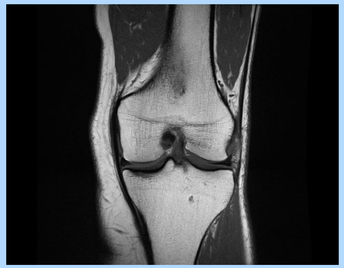
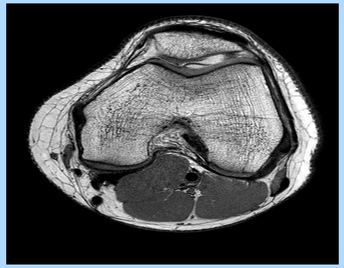
Navigating Challenges in AI-Assisted Knee MRI Reporting: A Case-Based Study

Dr Kasim Alimohamed, Dr Sabeeh Syed, Dr Ravindran Karthigan, Dr Samiullah Dost
Guy's and St Thomas' NHS Foundation Trust

Guy's and St Thomas' NHS Foundation Trust







CONTACT

Dr Kasim Alimohamed
Guy's and St Thomas NHS Foundation Trust
Email: Kasim.Alimohamed@nhs.net
Website:
http://www.guysandstthomas.nhs.uk

INTRODUCTION

Artificial Intelligence (AI) is transforming radiology by enhancing diagnostic workflows, improving accuracy, and minimizing human error[1]. AI as a potential solution is of increasing importance as the healthcare landscape transitions into an era of unprecedented demand, with tighter strains on resources[2]. This study evaluated the performance of a common large language model (LLM) in generating conclusions for knee MRI reports and assessed its accuracy compared to consultant musculoskeletal (MSK) radiologists.

METHODS AND MATERIALS

36 MRI knee reports were analysed. Using an AI large language model (gpt-3.5-turbo), conclusions were generated for these reports, which were compared against radiologists' conclusions. Accuracy was rated on a scale of 1 to 10. Statistical analysis was performed to calculate the mean accuracy score, standard deviation (SD), and 95% confidence interval (CI).

RESULTS

The mean accuracy score for AI-generated conclusions was 6.64 (SD = 2.29), with a 95% CI of 5.86 to 7.42, indicating moderate agreement with radiologist conclusions. The variability in scores highlighted instances where the AI struggled to interpret complex findings, particularly in nuanced diagnoses.

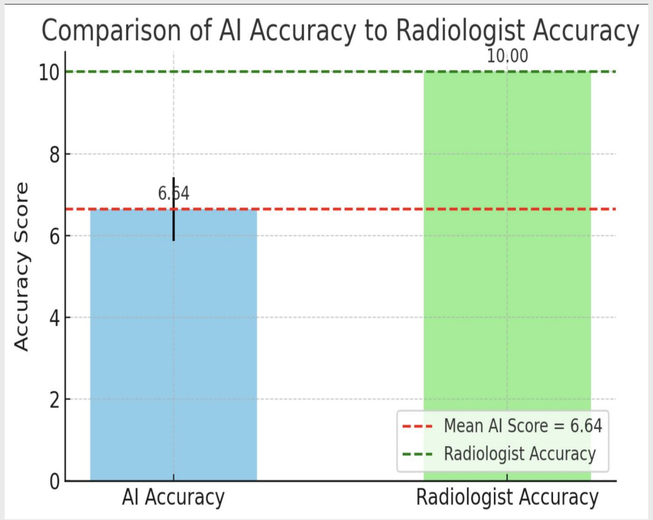


Figure 1: AI Accuracy vs Radiologist Accuracy

CONCLUSION

This study demonstrates AI's potential in generating clinically relevant summaries for knee MRI reports, improving efficiency and consistency in routine cases. However, limitations were observed in complex cases requiring detailed clinical judgment, especially when subtle or context-dependent findings were involved. This aligns with existing literature, which recognises both AI's strengths in streamlining workflows and its challenges in handling more nuanced diagnostic scenarios. To maximise AI's potential, further refinement of algorithms is needed to improve performance in complex cases and incorporate patient-specific data. AI should complement, not replace, the expertise of radiologists. A collaborative approach, combining AI with human judgment, is essential for optimising diagnostic accuracy and patient care.

REFERENCES

1. Ranschaert, E., Topff, L. and Pianykh, O. (2021) 'Optimization of radiology workflow with Artificial Intelligence', *Radiologic Clinics of North America*, 59(6), pp. 955–966. doi:10.1016/j.rcl.2021.06.006.
2. Sneag, D.B. *et al.* (2023) 'MRI advancements in musculoskeletal clinical and research practice', *Radiology*, 308(2). doi:10.1148/radiol.230531.