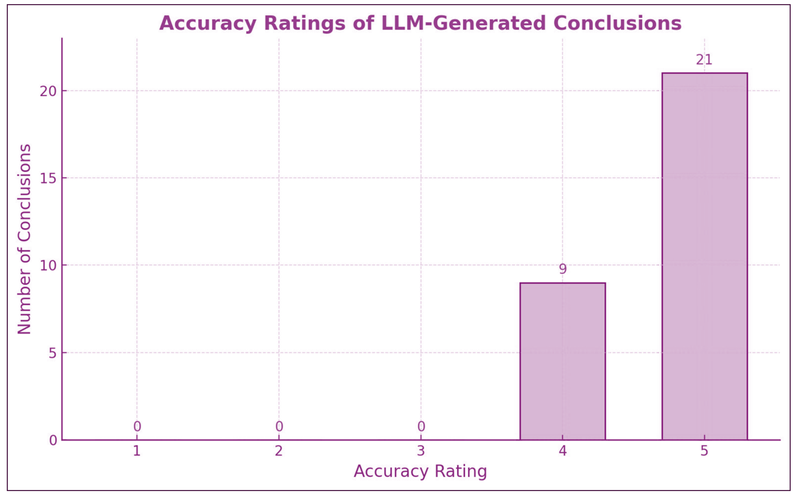
AI-Assisted Spine MRI Reporting: A Correlation Study with Radiologist-Derived Conclusions

Dr Avesta Ala'Aldeen, Dr Sabeeh Sayed, Dr Samiullah Dost
Department of Radiology, Guy's and St. Thomas' NHS Foundation Trust

Purpose

To evaluate the accuracy of a large language model (LLM) in generating conclusions for spine MRI reports with acute abnormal findings, as compared to radiologist-derived conclusions

This study serves as a proof of concept to assess the potential for LLMs to assist in radiology workflows and improve efficiency.



Methods and Materials

A total of 30 spine MRI reports with acute abnormal findings were selected from a tertiary centre. The report conclusions were generated independently by an LLM and then compared to the original radiologist-derived conclusions.

Each conclusion was rated on a scale of 1 to 5 based on its alignment with key findings and clinical significance:

5: Completely matches key findings and clinical significance
4: Captures most critical observations
3: Partially captures main points
2: Misses significant clinical implications
1: Substantially different or missed critical findings

Statistical analysis was conducted to calculate the weighted average rating and evaluate the distribution of accuracy scores using a chi-square goodness-of-fit test.

Results

The majority of LLM-generated conclusions achieved high accuracy ratings, with 21 conclusions rated as 5 and 9 rated as 4. The weighted average score was 4.7, reflecting strong agreement with radiologist-derived conclusions.

The chi-square test demonstrated a statistically significant deviation from a uniform distribution, indicating a strong concentration of high-accuracy scores.

Conclusion

This study demonstrates that LLMs can generate accurate conclusions for spine MRI reports, closely aligning with radiologist-derived conclusions.

The findings highlight the potential of LLMs to streamline radiology workflows by providing efficient and accurate preliminary summaries. Further studies are needed to refine these models and integrate them into clinical practice.

NHS
Guy's and St Thomas'
NHS Foundation Trust