View as "PDF version"

Patient Safety Tip of the Week

July 22, 2025

What the Heck are LLM's?

We've recently seen a lot about LLM's in the medical literature. So, what the heck are they? LLM's are **large language models**, an artificial intelligence concept. They have now been successfully used to identify errors in texts such as radiology reports.

Errors are actually fairly common in radiology reports. Error rates in radiology reports have been reported as approximately 3-5% in daily practice, even reaching around 30% when focusing on studies with abnormal findings (Brady 2017). Many radiologists are dictating their reports and errors can occur during transcription due to accents, background noise, interruptions, use of unfamiliar terms, technical issues, etc. Such errors are also more likely to occur during periods of excessive workload.

Radiology errors can include underreading, overreading, faulty reasoning, abnormalities outside area of interest (but visible), failure to consult prior imaging, inaccurate or incomplete clinical history, correct report failing to reach appropriate physician, and failure to see prior error leading to perpetuation of error through consecutive studies. So, when we are talking about radiology report errors or discrepancies, we are not just talking about typos or spelling errors! We could be talking about errors that have significant clinical implications.

Sun et al. (Sun 2025) constructed a dataset comprising 1656 synthetic reports (828 error-free, 828 with errors) and 614 reports from a chest radiograph dataset (307 error-free, 307 reports with errors). Examples of the errors included negation errors, interval change errors, left/right errors, and transcription errors (transcription errors were the largest category). They then applied their "fine-tuned LLM model" to basically proofread all those reports. You'll have to go to the full Sun paper for technical details of the LLM model. Reported errors were reviewed by a board-certified radiologist and a radiology resident. Their model demonstrated high accuracy in detecting errors, with 99 of 200 model identified reports confirmed by both radiologists and 163 confirmed by at least one radiologist.

Another study (Bhayana 2025) looked to see whether LLM's could improve reporting of O-RADS scores. O-RADS, the Ovarian-Adnexal Reporting and Data System for MRI, helps assign malignancy risk but adoption by radiologists has been inconsistent. The authors applied 2 LLM-based strategies to 284 pelvic MRI reports. They found that a hybrid LLM-based application, combining LLM feature classification with "deterministic" elements, accurately assigned O-RADS MRI scores from report descriptions, exceeding both an LLM-only strategy and the original reporting radiologist.

Another study (Mathai 2025) used an LLM to track interval changes in longitudinal radiology reports. The authors found their LLM for matching longitudinal report findings and tracking interval changes showed moderate to substantial agreement with the reference standard. For structured reporting, the LLM can pre-fill the "Findings" section of the next follow-up exam report with a summary of longitudinal changes to important findings. It can also enhance communication between the referring physician and radiologist.

Two other recent articles discuss the potential uses, unanswered questions, and research opportunities for LLM's in radiology (Marrocchio 2025, Bluethgen 2025).

Using artificial intelligence to read other reports has had other successes. Our August 29, 2023 Patient Safety Tip of the Week "The Perils of Copy and Paste" discussed the dangers of copy/paste. In that column we noted a study from China (Cheng 2022) that successfully restricted use of copy and paste in the EMR by using artificial intelligence (AI). They used electronic tools to detect word template similarities between clinical notes to identify those copied and pasted from previous visit notes. After developing an institutional policy, they set a threshold for determining whether a progress note was copied and pasted at 70% similarity to previous documents (using natural language programming and text mining) to determine the similarity. If the similarity was more than 70%, the computer would not save the progress note, similar to a plagiarism detection checker. The prevalence of copying and pasting was significantly reduced, from 35.7% to 23.7% (P=.001), after monitoring. The prevalence of copying and pasting initially showed a decreasing trend for 11 months, followed by a short period of a significantly increasing trend and then stability after the restriction of copying and pasting.

We concluded that column by saying we'll leave it up to the IT experts to figure out ways we can use natural language programming, text mining, and artificial intelligence to reduce overall use of copy & paste in the EMR.

While artificial intelligence has tremendous potential in healthcare, practical applications have been relatively scarce to date. Of course, AI also has the danger of using inaccurate or flawed data to come to erroneous conclusions. But the above examples show that AI research, particularly that using LLM's, seems to be on the right track to create solutions that improve clinical care and patient safety.

Some of our prior columns on patient safety issues in the radiology suite:

- October 16, 2007 "Radiology as a Site at High-Risk for Medication Errors"
- February 19, 2008 "MRI Safety"
- September 16, 2008 "More on Radiology as a High Risk Area"
- October 7, 2008 "Lessons from Falls....from Rehab Medicine"
- October 2008 "Preventing Infection in MRI"
- March 17, 2009 "More on MRI Safety"
- March 2009 "Risk of Burns during MRI Scans from Transdermal Drug Patches"
- August 11, 2009 "The Radiology Suite…Again!"
- January 2010 "Falls in the Radiology Suite"
- August 2010 "Sedation Costs for Pediatric MRI"
- January 25, 2011 "Procedural Sedation in Children"
- February 1, 2011 "MRI Safety Audit"
- October 25, 2011 "Renewed Focus on MRI Safety"
- March 13, 2012 "Medical Emergency Team Calls to Radiology"
- August 2012 "Newest MRI Hazard: Ingested Magnets"
- October 22, 2013 "How Safe Is Your Radiology Suite?"
- February 25, 2014 "Joint Commission Revised Diagnostic Imaging Requirements"
- July 2014 "New MRI Risks: for Staff!"
- July 1, 2014 "Interruptions and Radiologists"
- November 2014 "More Radiologist Interruptions"
- October 21, 2014 "The Fire Department and Your Hospital"
- June 23, 2015 "Again! Mistaking Antiseptic Solution for Radiographic Contrast"
- August 25, 2015 "Checklist for Intrahospital Transport"
- March 22, 2016 "Radiology Communication Errors May Surprise You"
- August 2016 "Guideline Update for Pediatric Sedation"
- October 2016 "MRI Safety: There's an App for That!"
- January 17, ,2017 "Pediatric MRI Safety"
- August 8, 2017 "Sedation for Pediatric MRI Rising"
- November 14, 2017 "Tracking C. diff to a CT Scanner"
- March 2018 "MRI Death a Reminder of Dangers"
- March 2018 "Cardiac Devices Safe During MRI But Spinners!?"
- April 2018 "Radiologists Get Fatigued, Too"
- May 2018 "Cost of Interrupting a Radiologist"
- November 2018 "OMG! Not My iPhone!"
- December 11, 2018 "Another NMBA Accident"
- April 2, 2019 "Unexpected Events During MRI"
- September 2019 "New MRI Hazard: Magnetic Eyelashes"
- October 15, 2019 "Lots More on MRI Safety"
- November 5, 2019 "A Near-Fatal MRI Incident"
- November 12, 2019 "Patient Photographs Again Help Radiologists"
- November 26, 2019 "Pennsylvania Law on Notifying Patients of Test Results"

- January 7, 2020 "Even More Concerns About MRI Safety"
- July 14, 2020 "A Thesis on Intrahospital Transports"
- August 25, 2020 "The Off-Hours Effect in Radiology"
- October 2020 "New Warnings on Implants and MRI"
- January 2021 "New MRI Risk: Face Masks"
- May 25, 2021 "Yes, Radiologists Have Handoffs, Too"
- June 1, 2021 "Stronger Magnets, More MRI Safety Concerns"
- November 2021 "Yet Another Risk During MRI"
- January 2022 "MRI Safety Issues"
- February 2022 "Managing Incidental Findings"
- June 2022 "Where Are You Barcoding?"
- July 12, 2022 "Radiologists Racked by Interruptions"
- July 26, 2022 "More Risks in the Radiology Suite"
- May 2023 "Factors Associated with Radiology Diagnostic Errors"
- September 12, 2023 "Radiology and Hospital-Acquired Infections"
- October 24, 2012 "Serious MRI Accident at Unregulated MRI Center"
- November 21, 2023 "Another Terrifying MRI Accident"
- January 2024 "Guns and MRI Don't Mix"
- June 11, 2024 "Nursing in the MRI Suite"
- August 13, 2024 "Preventing Thermal Injury During MRI"
- October 2024 "Another MRI Accident Reveals Gap in Oversight
- October 2024 "MRI Special Patient Considerations"
- October 8, 2024 "First Responders and MRI"
- April 22, 2025 "Updated ACR Manual on MRI Safety"
- July 8, 2025 "Applying Human Factors Concepts to Radiology"

References:

Brady AP. Error and discrepancy in radiology: inevitable or avoidable? Insights Imaging 2017; 8(1): 171-182
https://pmc.ncbi.nlm.nih.gov/articles/PMC5265198/

Sun C, Teichman K, Zhou Y, et al. Generative Large Language Models Trained for Detecting Errors in Radiology Reports. Radiology 2025; 315: 2
https://pubs.rsna.org/doi/10.1148/radiol.242575

Bhayana R, Jajodia A, Chawla T, et al. Accuracy of Large Language Model–based Automatic Calculation of Ovarian-Adnexal Reporting and Data System MRI Scores from Pelvic MRI Reports. Radiology 2025; 315: 1
https://pubs.rsna.org/doi/10.1148/radiol.241554

Mathai, T.S., Kim, B., Stroie, O.M. et al. Privacy-Preserving Large Language Model for Matching Findings and Tracking Interval Changes in Longitudinal Radiology Reports. J Digit Imaging Inform Med 2025; Published April 11, 2025
https://link.springer.com/article/10.1007/s10278-025-01478-7#citeas

Marrocchio C, Sverzellati N. Will Generative Large Language Models Become Radiologists' Invaluable Allies? Radiology 2025; 315: 2
https://pubs.rsna.org/doi/10.1148/radiol.251259

Bluethgen C, Van Veen D, Zakka C, et al. Best Practices for Large Language Models in Radiology. Radiology 2025; 315: 1
https://pubs.rsna.org/doi/10.1148/radiol.240528

Cheng CG, Wu DC, Lu JC, Yu CP, Lin HL, Wang MC, Cheng CA. Restricted use of copy and paste in electronic health records potentially improves healthcare quality. Medicine (Baltimore). 2022; 101(4): e28644
https://pmc.ncbi.nlm.nih.gov/articles/PMC8797538/

The Truax Group
Healthcare Consulting
www.patientsafetysolutions.com

http://www.patientsafetysolutions.com/

Home

Tip of the Week Archive

What's New in the Patient Safety World Archive