

AI-driven Subspecialist Allocation for CT Body Examinations: Impact on Report Quality

Dr Jaymin Patel, Dr Adam Rushford, Dr Robert Dunk, Dr Rebecca Greenhalgh
Hexarad Group, London, United Kingdom



Background

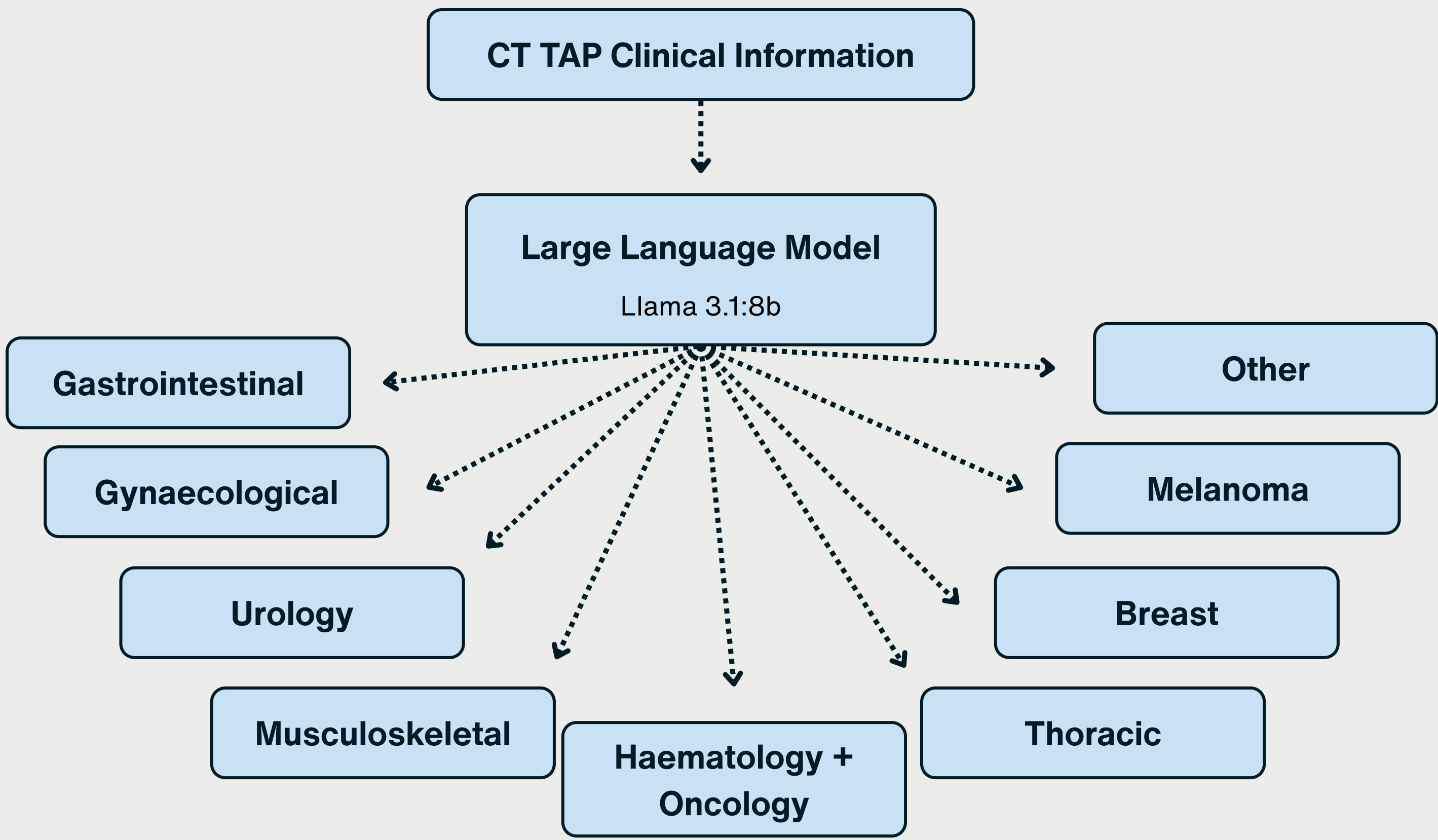
CT chest, abdomen, and pelvis (CT TAP) examinations are typically reported by general or body radiologists because, unlike named protocols such as CT pancreas, the examination title alone does not indicate the relevant subspecialty. The clinical information required to determine optimal allocation – for example, to a gastrointestinal, gynaecological, or chest radiologist – is buried within the free-text referral, often abbreviated or ambiguous. This is a global challenge, particularly in smaller general hospitals and teleradiology settings¹. As a result, CT TAP studies are rarely triaged to subspecialists, despite evidence that subspecialist reporting improves diagnostic accuracy². We developed an in-house large language model (LLM) tool to parse free-text clinical information on CT TAP requests and categorise each into a subspecialty, enabling automatic allocation to the appropriate subspecialist. We aimed to evaluate its impact on report quality as measured by peer review outcomes.

Methods

We conducted a retrospective before-and-after analysis of peer review data for CT TAP examinations across our teleradiology network. The pre-intervention period preceded deployment of the AI allocation tool; the post-intervention period followed full deployment. Primary outcomes were the proportion of reports achieving full peer-review agreement (scored on a 1-6 ordinal scale) and the addendum rate. The control group consisted of non-CT TAP peer review data over the same periods (n=8,466 pre-intervention; n=10,787 post-intervention.)



LLM Allocation Process



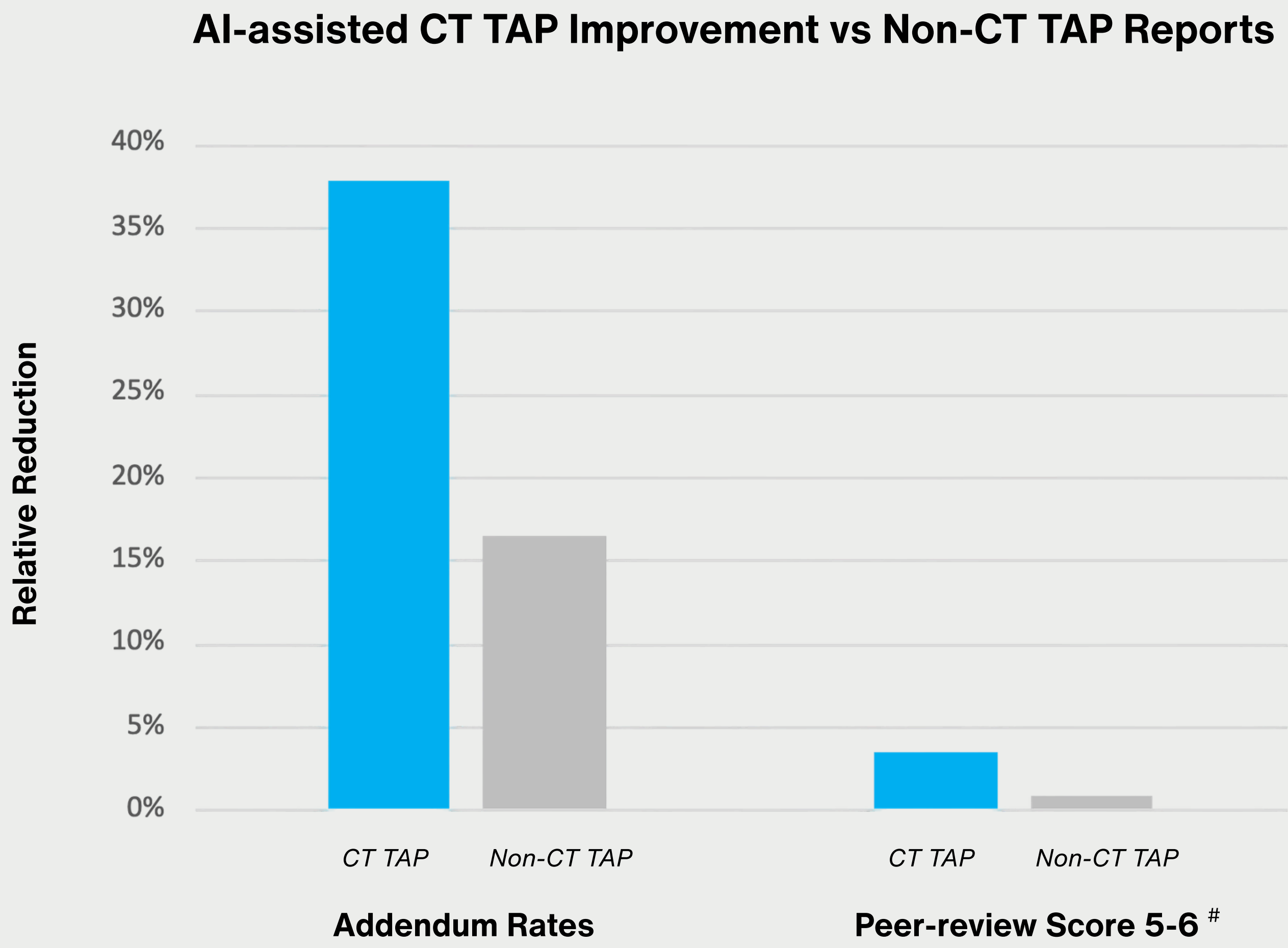
Outcome Definitions

- **Addendum Rate:** Proportion of peer-reviewed reports requiring a clinically meaningful amendment after finalisation.
- **Peer-Review Agreement:** Proportion of peer-reviews that received the highest scores (5-6 on a 6-point ordinal scale), indicating no clinically significant discrepancy between the original and reviewing reporter.

Results

37.8% Reduction in Addendum Rates

- A total of 988 CT TAP peer reviews were analysed pre-intervention and 1,459 post-intervention, following deployment of the AI-driven subspecialist allocation tool
- The addendum rate among CT TAP reports fell by 37.8% relatively in the post-intervention period, a statistically significant reduction ($p<0.001$)
- Full agreement (Peer-review score 5-6) rose from 86.84% to 89.92% ($p=0.02$)



A peer-review score of 5 or a 6 represents no discrepancy or a 'great spot'

Across all non-CT TAP reports, full agreement rose 0.79% and the addendum rate fell 16.5% relatively. Improvement in the AI-assisted CT TAP subset was approximately four times greater than the all-modality baseline on both measures.

Conclusion

- LLM-driven subspecialist allocation of CT TAP referrals produced a clinically meaningful improvement in report quality
- The magnitude of improvement substantially exceeded background organisational trends, supporting attribution to the intervention
- The LLM-based system addresses the difficulty in manual subspecialty allocation, suggesting this may be a high-value, low-risk application of AI in radiology

Further Research

Addendum rates and peer-review scores reflect a blinded marker of report quality. Further research incorporating feedback from referring clinicians and MDT-lead radiologists would provide a comprehensive assessment of the intervention's downstream impact on patient care.

1. Wu et al. Clin Radiol. 2020, 75(11), 879.e7–879.e11.
2. Siewart et al. JACR. 2025, 22(1), 129-138

Contact

