

DIAGNOSTIC ACCURACY OF AN AI-DRIVEN PRIORITISATION TOOL FOR THE IDENTIFICATION OF HIGH-CONFIDENCE NORMAL CT HEAD STUDIES

Jack Sheppard¹, Stefan T. Schwarz^{1,2}, Hisham Jaber^{1,3}, Christopher Goodwin^{1,3}, Mel Ryan⁴, Alicja Raginis-Zborowska⁴, Thomas Pickersgill⁴, Caroline Parker¹, Dane Evans¹, Phillip Wardle^{1,3}, Kieran G. Foley^{1,3}



Academi Ddelweddu
Genedlaethol Cymru
National Imaging
Academy Wales

¹ National Imaging Academy Wales (NIAW) ² Cardiff & Vale University Health Board ³ Cwm Taf Morgannwg University Health Board ⁴ Harrison.ai

PURPOSE

Radiology departments are struggling to meet the increasing demand for emergency CT head imaging. Many rely on costly teleradiology companies to provide timely reports. We evaluated the utility of an AI-prioritisation tool to identify high-confidence clinically normal non-contrast CT head studies. The primary aim was to test the agreement between the AI-prioritisation tool and reference index for categorisation. The secondary aim was to determine the standalone diagnostic accuracy (sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV)) of the AI-prioritisation tool for categorisation of “Urgent” studies.

METHODS AND MATERIALS

Consecutive CT studies performed out-of-hours in two University Health Boards (UHBs) in July 2025 were included in this retrospective study. Images and reports were anonymised and transferred to a secure PACS where an AI tool inferred all studies. The AI-prioritisation tool categorised each CT as “Urgent” or “Not Urgent” based on the clinical acuity of findings. A consultant neuroradiologist reviewed each CT providing the reference index classification (“Urgent” or “Not Urgent”). All studies were originally reported by either a resident radiologist or teleradiologist. The consultant neuroradiologist assigned their classification before reviewing the original report. “Urgent” reports were defined as any CT with positive findings that may have influenced clinical management out-of-hours. The primary outcome was the agreement between the AI-prioritisation tool and consultant neuroradiologist, defined as the proportion of cases correctly categorised. The study abstract was written in accordance with the STARD abstract guidelines for diagnostic accuracy studies.

SELECTED CASES: TRUE POSITIVES – CASES APPROPRIATELY PRIORITISED AS URGENT

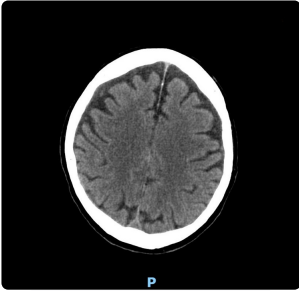


FIGURE 1: BILATERAL SUBDURAL HYGROMAS – A “GOOD SPOT”, NOT REPORTED BY REPORTING RADIOLOGIST BUT PICKED UP BY AI TOOL.

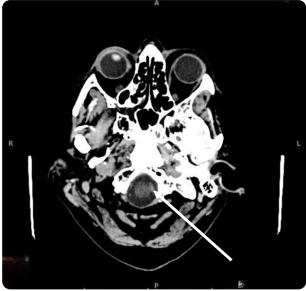


FIGURE 2: LEFT VERTEBRAL ARTERY (LVA) CLOT AND CALCIFIED CHOROID PLEXUS – URGENT PRIORITIZATION BY AI (FOR COLLOID CYST (ACTUALLY CHOROID PLEXUS)) AND ALSO BY THE CONSULTANT NEURORADIOLOGIST (LVA CLOT).

RESULTS

615 CT head examinations were included (UHB1=443, UHB2=172). 166/615 (27.0%) were prioritised as “Urgent” by the AI-prioritisation tool. 94/615 (15.2%) were prioritised as “Urgent” by the consultant neuroradiologist. Agreement was 86% (95% CI [83, 88%]). Sensitivity, specificity, PPV, and NPV of the standalone diagnostic tool for classification of “Urgent” studies were 91% [84-96%], 85% [81-88%], 52% [44-60%], and 98% [97-99%], respectively. Correct categorisation of “Urgent” by the AI-prioritisation tool, but for the incorrect reason, was demonstrated in two cases (0.3%) (Figure 2).

SELECTED CASES: FALSE POSITIVES – CASES INAPPROPRIATELY PRIORITISED AS URGENT

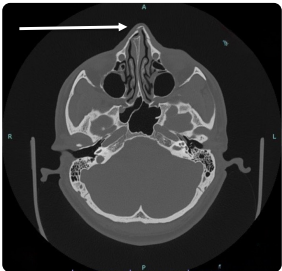


FIGURE 3: NASAL BONE FRACTURE – CORRECTLY IDENTIFIED BY AI TOOL BUT INAPPROPRIATELY PRIORITISED AS URGENT. NOT URGENT PRIORITISATION BY CONSULTANT NEURORADIOLOGIST.

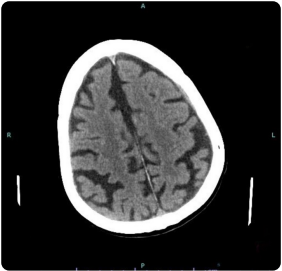


FIGURE 4: LEFT-SIDED EXTRA-AXIAL HAEMATOMA, IMPROVING COMPARED TO PREVIOUS – CORRECTLY IDENTIFIED BY AI TOOL BUT INAPPROPRIATELY PRIORITISED AS URGENT. NOT URGENT PRIORITISATION BY CONSULTANT NEURORADIOLOGIST.

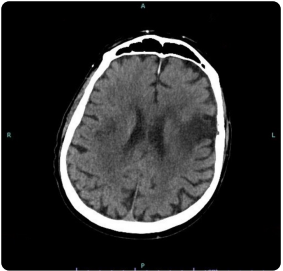


FIGURE 5: CHRONIC LEFT-SIDED INFARCT – CORRECTLY IDENTIFIED BY AI TOOL BUT INAPPROPRIATELY PRIORITISED AS URGENT. NOT URGENT PRIORITISATION BY CONSULTANT NEURORADIOLOGIST.



FIGURE 6: SEPTUM PELLUCIDUM – INCORRECTLY LABELLED HYDROCEPHALUS AND PRIORITIZED AS URGENT BY THE AI TOOL. NOT URGENT PRIORITISATION BY CONSULTANT NEURORADIOLOGIST.

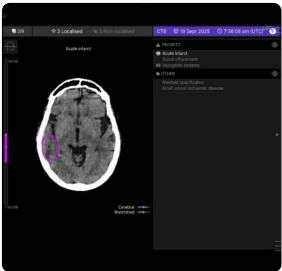


FIGURE 7: ARTEFACTUAL APPEARANCES DESIGNATED AS ACUTE INFARCT – INCORRECT URGENT PRIORITIZATION BY AI TOOL. NOT URGENT PRIORITISATION BY CONSULTANT NEURORADIOLOGIST.

SELECTED CASES: FALSE NEGATIVES – CASES INAPPROPRIATELY PRIORITISED AS NON-URGENT

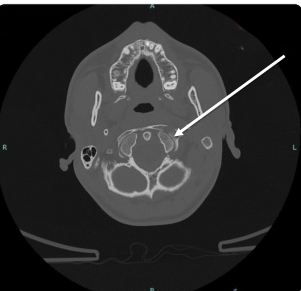


FIGURE 8: UNSTABLE C2 FRACTURE – MISSED BY AI TOOL AND PRIORITIZED AS NOT URGENT. URGENT PRIORITISATION BY CONSULTANT NEURORADIOLOGIST.

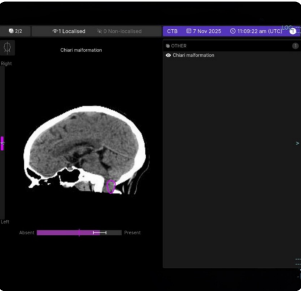


FIGURE 9: CHIARI I MALFORMATION – CORRECTLY IDENTIFIED BY AI TOOL BUT PRIORITISED AS NOT URGENT. URGENT PRIORITISATION BY CONSULTANT NEURORADIOLOGIST.

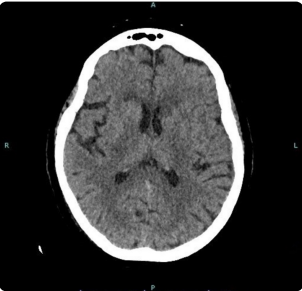


FIGURE 10: ACUTE INFARCT OF THE LEFT CAUDATE NUCLEUS – MISSED BY AI TOOL AND PRIORITISED AS NOT URGENT. URGENT PRIORITISATION BY CONSULTANT NEURORADIOLOGIST.

CONCLUSION

The AI-prioritisation tool demonstrated good agreement, sensitivity and NPV compared to consultant neuroradiologists. Further prospective validation is required, however, the AI-prioritisation tool showed the potential for safe identification of high-confidence normal CT head studies. Of the abnormalities in the false negative group, cervical spine fractures made up the greatest proportion. The AI-prioritisation tool may enable deferred reporting of cases prioritised as “Not Urgent”, potentially reducing costs associated with outsourcing, and acknowledging specific limitations.