

Dr Sonam Vadera, Dr Aditya U Kale, Dr Qasim Malik, Mr Robert McEwan, Professor Alastair Denniston, Dr Xiaoxuan Liu

svadera13@gmail.com

Background & Methods

- Amid ever-growing interest in the potential benefits of AI in healthcare, there are concerns around how devices should be monitored for performance and safety, with no standardised evaluation pathway.
- The Medical Algorithmic Audit (MAA)¹ is a safety monitoring framework for algorithmic error auditing and failure mode detection.
- We applied this framework to the evaluation of a deep learning AI system for autonomous reporting of normal chest radiographs.
- ChestLink, Oxipit, is a CE-marked Class IIb medical device which analyses chest X-rays in real-time. It either generates a "normal" report, bypassing radiologist review, or flags the study for further radiologist evaluation if pathologies cannot be ruled out with high confidence.

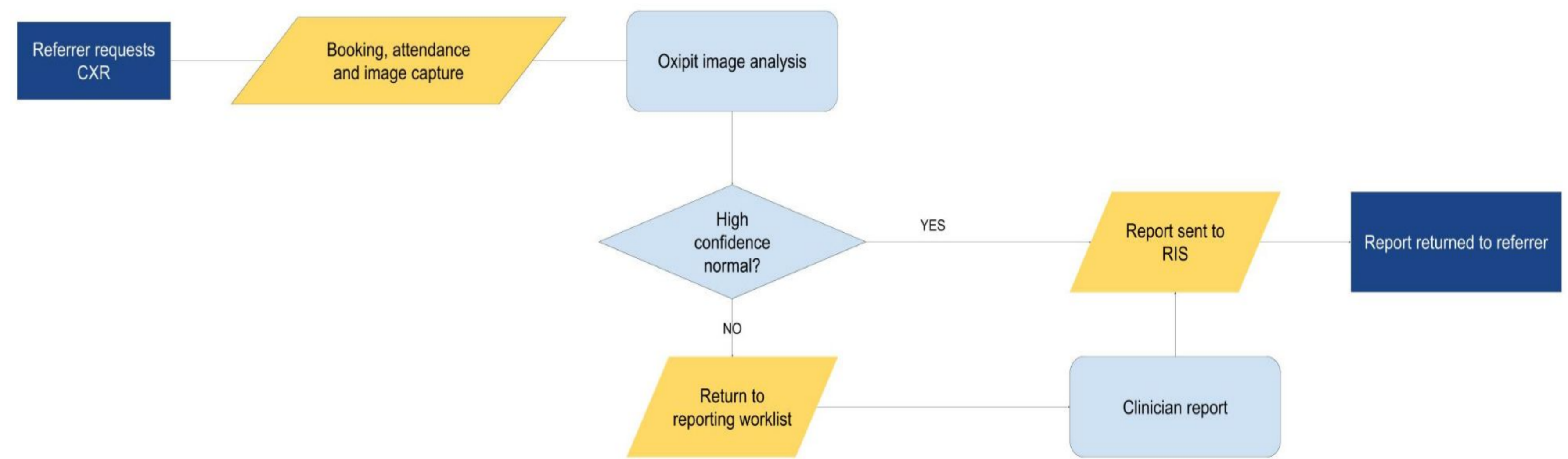


Figure 1: Flow chart demonstrating the AI-enabled pathway using Oxipit

Medical Algorithmic Audit framework as applied to Oxipit-enabled autonomous reporting

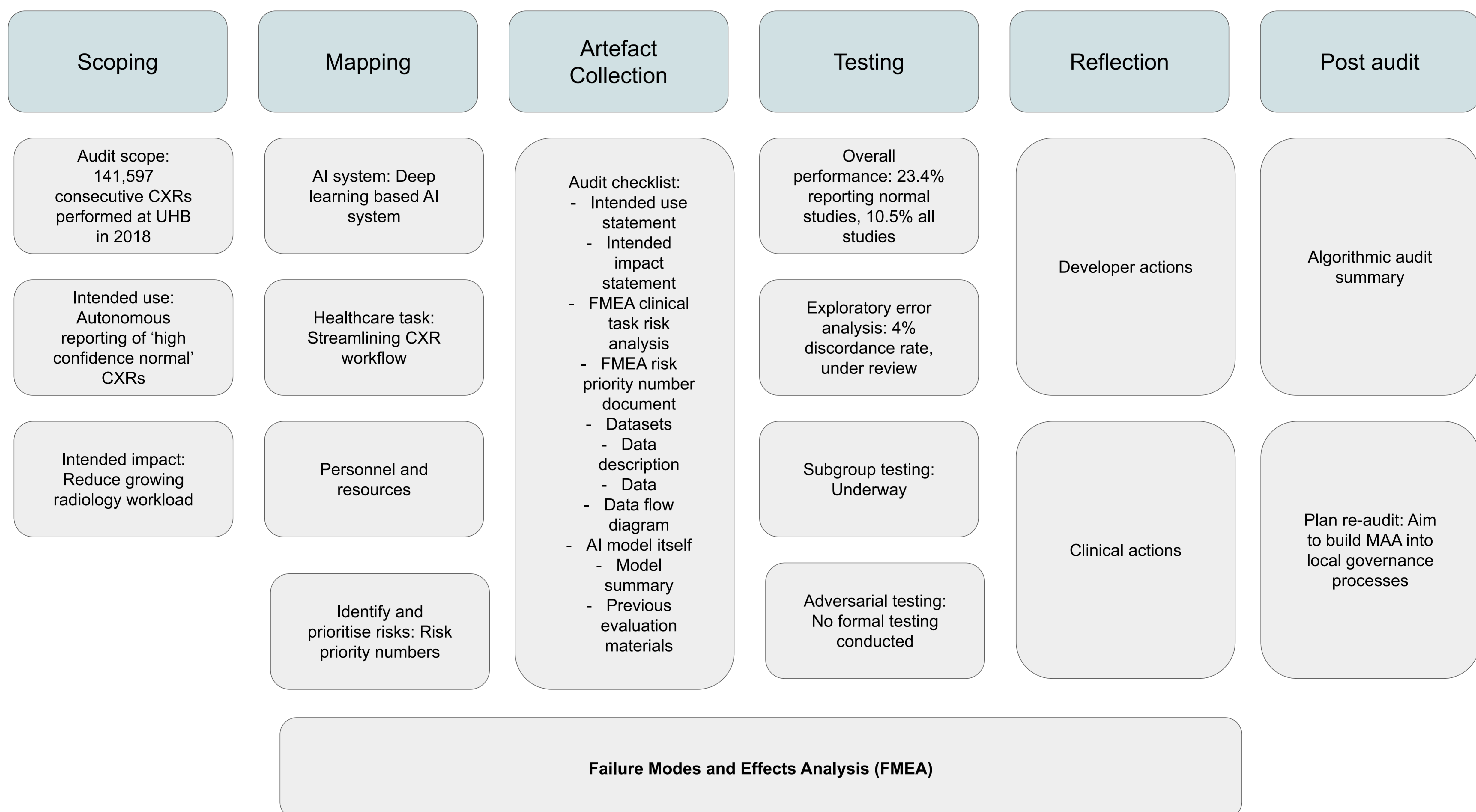


Figure 2: Overview of the MAA framework applied to ChestLink, Oxipit

Results

	All studies
In scope for ChestLink processing	141,597
Normal studies (<u>UHB report / NLP</u>)	63,442
Normal studies (<u>ChestLink</u>)	14,820
ChestLink reporting rate of <u>normals</u>	23.4%
ChestLink reporting rate of <u>all studies</u>	10.5%
Discordant cases (ChestLink normal & NLP abnormal)	593
Discordance rate	4.0%

- This audit examines 141,597 cases which were in scope for processing by Oxipit.
- ChestLink demonstrates promising accuracy in identifying normal chest X-rays, potentially reducing reporting workload. Results suggest ChestLink can report approximately ¼ of normal studies autonomously.
- Nonetheless, the observed discordance warrants further investigation. A case-by-case review by four consultant cardiothoracic radiologists is underway, and exploratory and subgroup error testing are also being conducted.

1. Liu, X. et al. (2022) 'The Medical Algorithmic Audit', *The Lancet Digital Health*, 4(5). doi:10.1016/s2589-7500(22)00003-6.

Discussion & Next Steps

- Safe deployment of AI requires robust monitoring. The MAA provides a standardised framework for understanding key risks using a collaborative, multistakeholder approach.
- The MAA for Oxipit is in testing phase to identify potential failure modes and inform the decision to proceed to the prospective trial stage and subsequent deployment.
- Further work will be required to understand the efficient use in practice, broader stakeholder perspectives and engagement, and optimal audit frequency.

Progress and Milestones

- Significant progress has been made on the MAA for Oxipit, and I have contributed to a recently completed MAA for DERM (Skin Analytics) tool for triage of skin lesions.
- Regular meetings with developers and UHB Digital Transformation Team. Discussions with legal teams and regulatory bodies around implications of using autonomous reporting devices in the NHS.
- Presentation of findings to key stakeholders.
- Other opportunities as part of wider UHB team: contribution to systematic review looking at dataset transparency/diversity, creation of mock target product profiles, experience in guiding SMEs at early stage of product development, attendance at meetings with international experts in the network.