

The use of Artificial Intelligence (AI) as a diagnostic aid in public medical imaging

Macy Lu and John Hillier

Abstract

The nature of radiology work means that we have broad opportunities to leverage on technology advances and AI to help us provide better patient care with the same workforce bottleneck. Beyond interpretation, AI is predicted to have a role throughout the entire workflow, from organizing imaging requisitions, to triaging protocols, acquiring images, improving image quality, and generating reports (Yan, et al., 2025).

A survey (Chen & Cui, 2025) found that distrust and aversion appear when patients know that the doctor uses ChatGPT, even if it serves as a supporting tool rather than a primary decision maker. This suggests that transparency, education and feedback is essential to build trust in AI technologies in health care. Health systems should create patient-centered AI governance processes and communications for a trustworthy ecosystem (Nong & Ji, 2025).

The purpose of this paper is to create awareness for shared learnings, in the setting of increasing use of AI in healthcare.

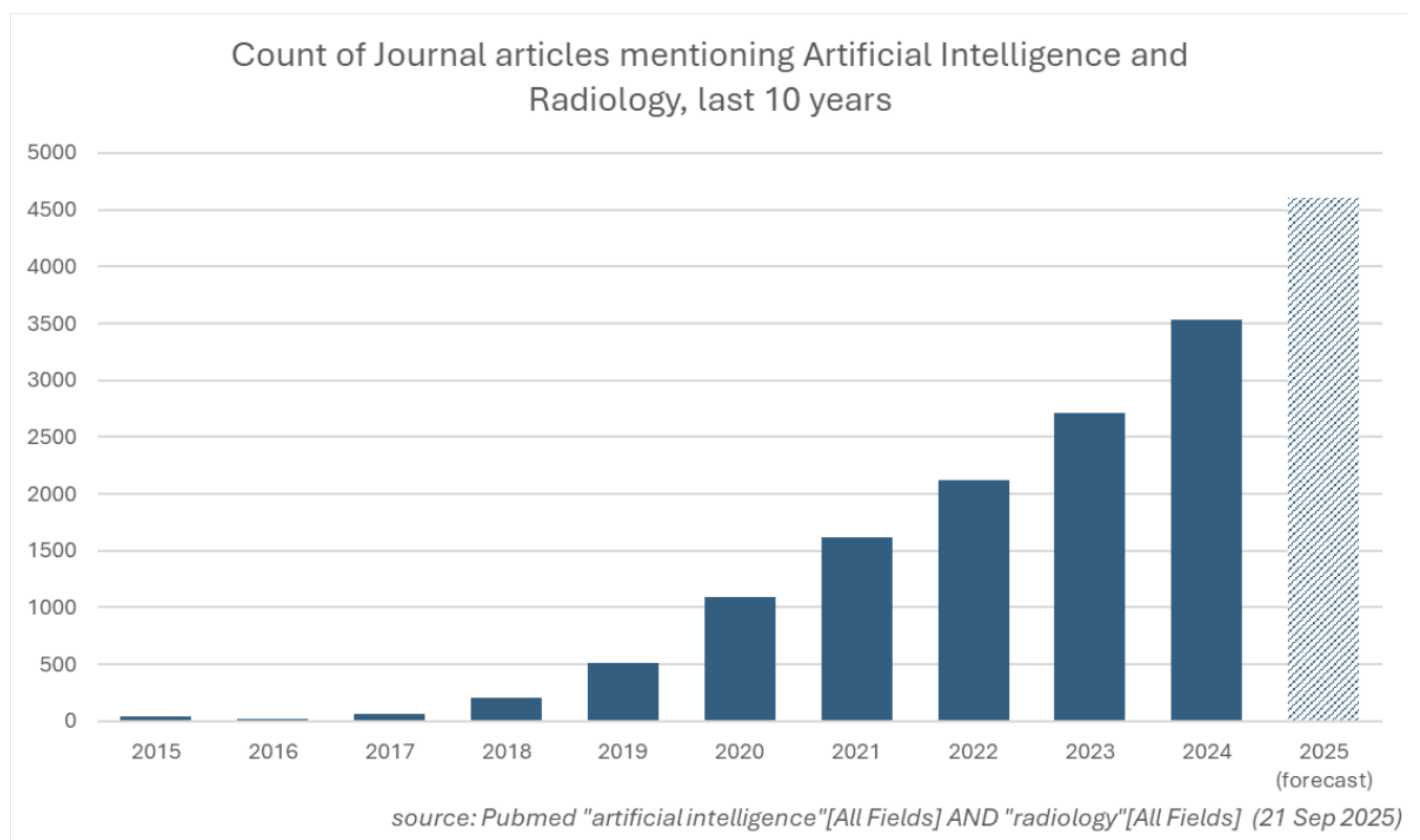


Figure 1: Created by John Hillier

78% of all cleared AI-enabled medical devices were for radiology in 2024 (Mousa, 2025).

AI pilot trial



Figure 2: 124 I C XR findings

feedback and find trends have led to a strong understanding of the system behaviours needed for clinical trust and implementation reliability. There has been close collaboration with the vendor (now Harrison.ai) who has been providing regular dashboard updates (eg. tool access volume, AI findings) and performing analysis into anecdotal feedback (eg. false positives or negatives).

Formal post-trial survey results are pending at this stage. Anecdotally, there are mixed reports of good picks (subtle true positives), false positives and false negatives by the radiologists, who are still building awareness of the benefits and risks of AI in this first dive into AI. On the clinician/referrer end, there is general appreciation of an early AI second opinion. Importantly, urgent findings on AI are now reported as priority to drive timely patient management.

Background

The primary bottleneck in Medical Imaging is workforce, including Radiologists - leading to a large unreported list. CXR AI can be used to triage urgent findings for priority reporting, diagnosis and treatment, improving service delivery and clinical outcomes. This cloud-based solution integrates with the Metro North Radiology Information Systems (RIS) Karisma and the Picture Archival Systems (PACS) IntelePACS. The secondary capture feature enables users with access to PACS visibility of AI generated images to allow point of care decision support (Figure 2).

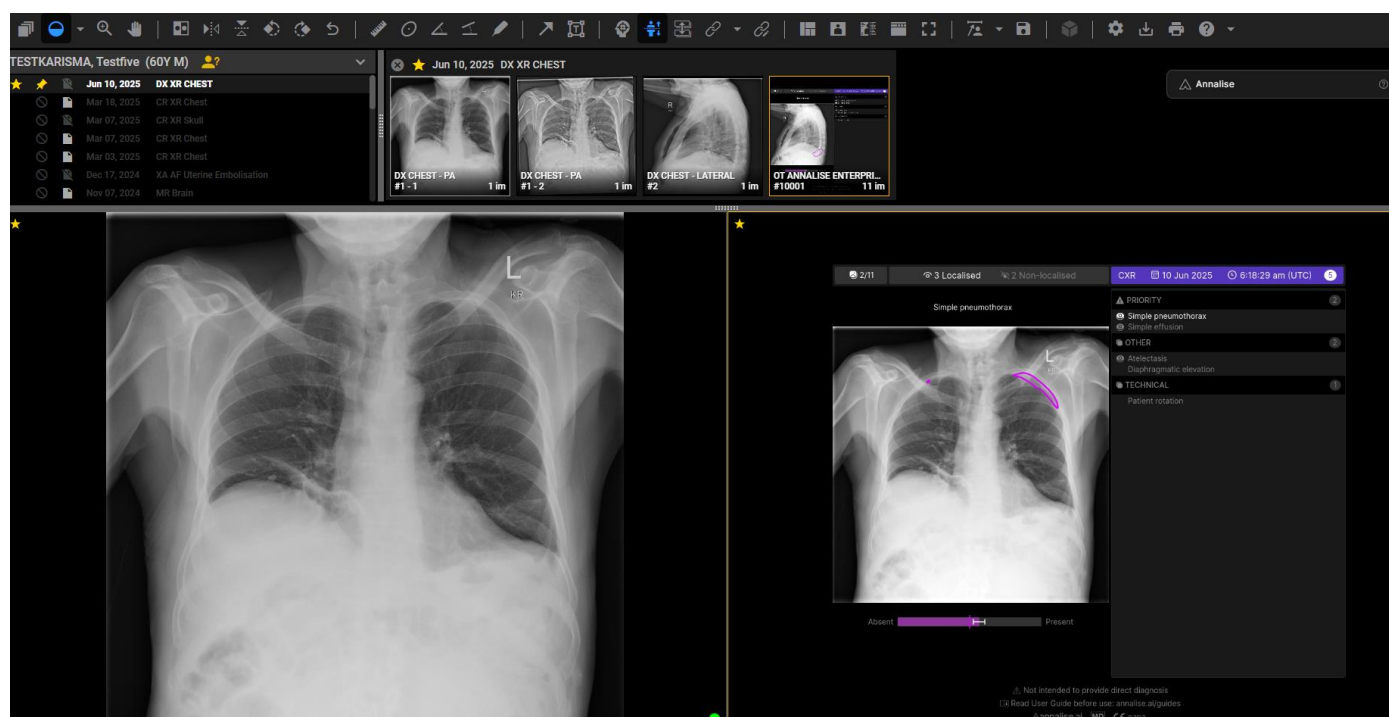


Figure 3: Example of left apical pneumothorax highlighted by AI

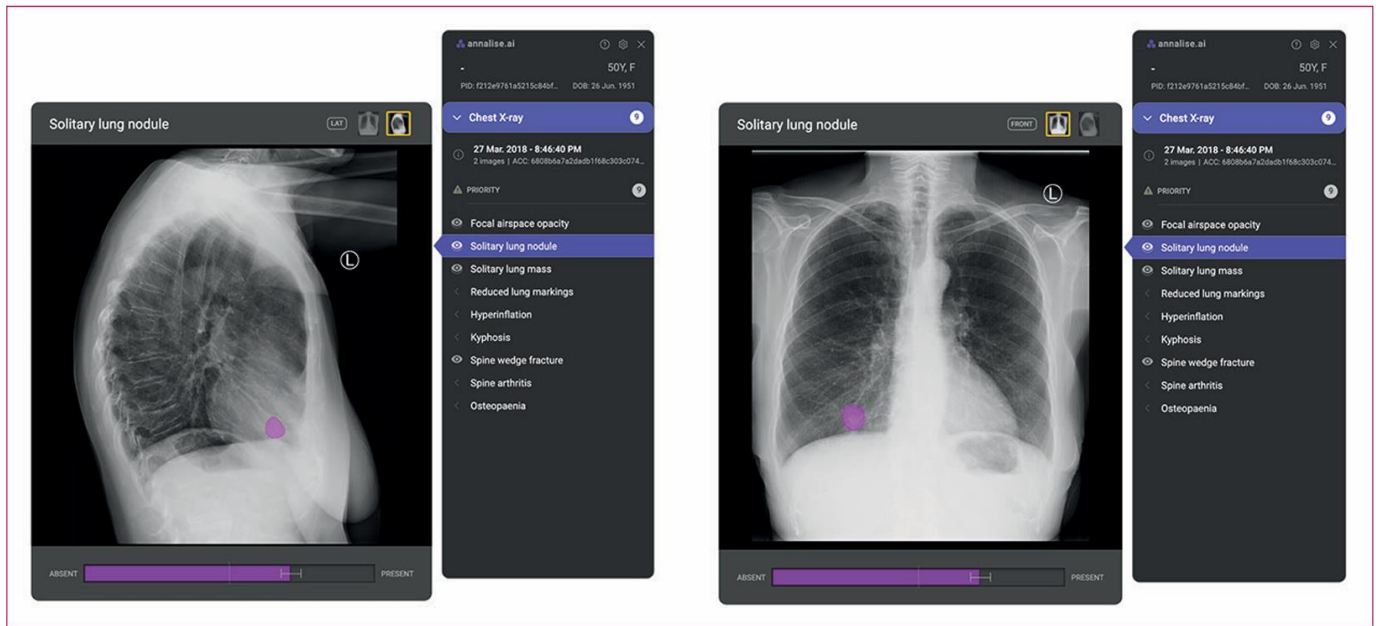


Figure 4: Example of right lung nodule highlighted by AI

Hence, education on the interpretation and use of these images must extend from the medical imaging department to the relevant hospital health care workers (Figures 3 and 4).

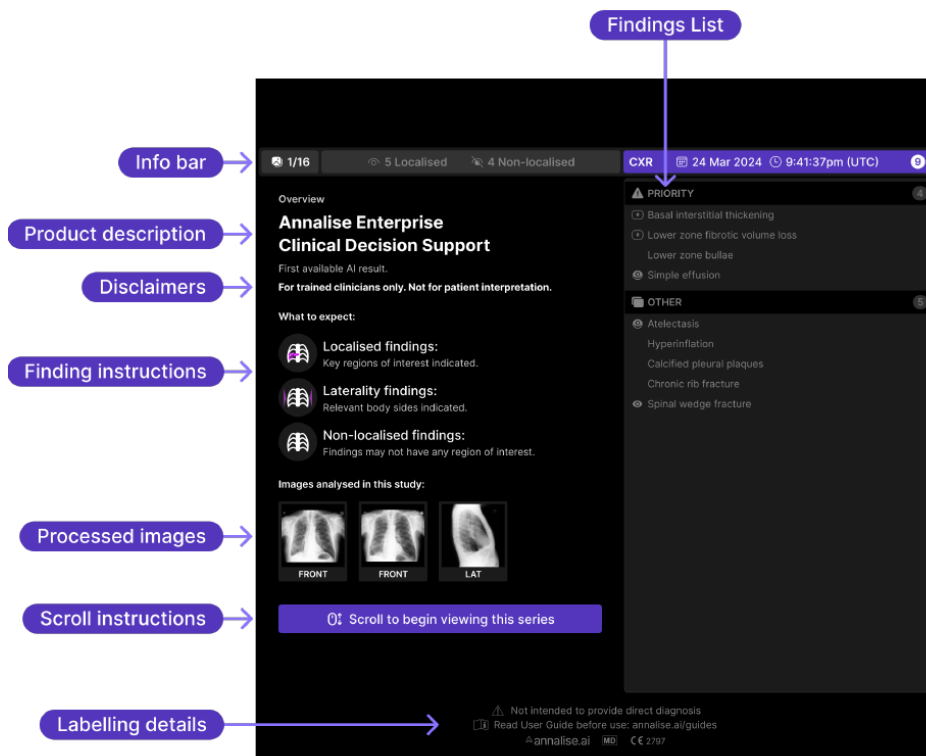


Figure 4: Explanation of AI use

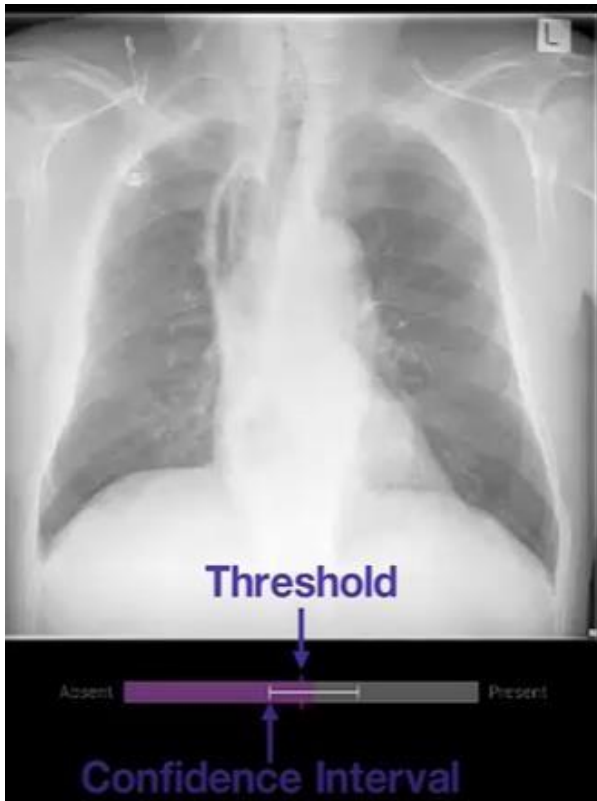


Figure 5: User understanding of the confidence reading to assess the likelihood of false versus true positives is essential.

It is crucial that all users are aware of the limitations of AI (Figure 5) to ensure appropriate utilisation; including the need for clinical and diagnostic acumen for determining false positive and negative AI results

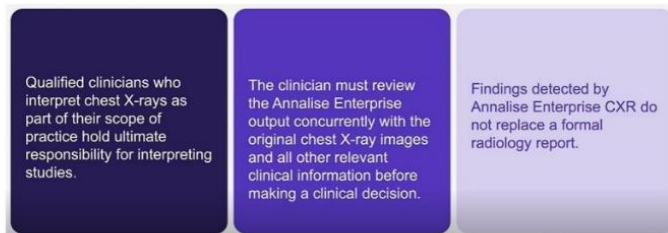


Annalise.AI – Considerations for MNHHS clinicians

Verify AI Findings – Always corroborate AI findings with clinical symptoms, patient history and formal radiology report (or if unavailable seek opinion from Medical Imaging) .

Use AI as a support tool – AI is not a substitute for clinical decision making or a substitute for the final radiology report. While AI can assist with clinical decision making, the final responsibility will always lie with the clinical team.

AI does not look at previous imaging – Each chest x-ray is interpreted in isolation with no reference or comparison to previous imaging. If available, always make comparison to previous imaging.



DMN



Annalise.AI – Limitations for MNHHS clinicians

AI is not to be used for confirmation of line or tube placement – The introduction of AI does not replace clinician review of images or clinical decision making for line/tube placement confirmation. This remains a clinical decision.

Requires due diligence when used solely to decide on Intervention (e.g. pleural drain) or CT follow-up – The decision to proceed with any interventional procedures, e.g. pleural aspiration/pleural drainage, or CT delineation, remains a clinical decision. Where there is ambiguity, clinicians should rely on clinical judgement, check formal radiology report (or if unavailable seek opinion from Medical Imaging), or further investigations before proceeding.

****If a clinician has any questions/concern with pathology findings from the AI generated chest x-ray image, they should contact their Senior Medical Officer or Radiologist on shift.**

DMN

Figure 6: Some specific considerations and limitations (not comprehensive) of the CXR AI tool

AI safety and principles

Artificial intelligence refers to an engineered system that generates predictive outputs (Department of Industry, Science and Resources, 2023) - taking actions at a speed and scale that would otherwise be impossible. The safe and responsible deployment and adoption of AI presents significant

opportunities for Australia to improve economic and social outcomes. Risks with all interpretation AI tools include false positive and false negative results, and inappropriate use by untrained individuals (including patients). The *Australian Community Attitudes to Privacy Survey 2020* prepared for the Office of the Australian Information Commissioner (OAIC) showed that 84% of respondents believed people should have a right to know if a decision affecting them is made using AI technology. 78% believed individuals should be told what factors and personal information are considered by the algorithm and how these factors are weighted. Australian commission on safety and quality in health care (2025) recommends that patients should be advised on how the AI tool supports care delivery. Hence, the following information may be useful for inclusion in all AI-supported CXR reports: “This report was aided by an AI decision-support tool, noting that not all AI findings may be deemed pertinent or accurate for inclusion into the report at the time of interpretation.”

There are other risks that are in-built into AI tools, such as algorithmic bias which involves systematic or repeated decisions that privilege one group over another. All AI risks must be considered and proactive steps taken to mitigate.

Australia's AI Ethics Principles (Department of Industry, Science and Resources, 2023):

1. **Human, societal and environmental wellbeing:** AI systems should benefit individuals, society and the environment.
2. **Human-centred values:** AI systems should respect human rights, diversity, and the autonomy of individuals.
3. **Fairness:** AI systems should be inclusive and accessible, and should not involve or result in unfair discrimination against individuals, communities or groups.
4. **Privacy protection and security:** AI systems should respect and uphold privacy rights and data protection, and ensure the security of data.
5. **Reliability and safety:** AI systems should reliably operate in accordance with their intended purpose.
6. **Transparency and explainability:** There should be transparency and responsible disclosure so people can understand when they are being significantly impacted by AI.
7. **Contestability:** When an AI system significantly impacts a person, community, group or environment, there should be a timely process to allow people to challenge the use or outcomes of the AI system.
8. **Accountability:** People responsible for the different phases of the AI system lifecycle should be identifiable and accountable for the outcomes of the AI systems, and human oversight of AI systems should be enabled.

The unintended consequences of AI such as deference to tools – eg. flagging of more mammogram findings that increases workup without improvement in cancer detection (Mousa, 2025), or over-reliance of radiology registrars during training - will need consideration.

Research

The trial commenced with a retrospective ground-truthing review by 6 radiologists of model performance to ascertain if the default diagnostic thresholds are satisfactory for the local patient population (Jones, C., 2025). A dataset of 286 chest radiographs over 3 consecutive weeks were assessed for a subset of 35/124 urgent findings. This demonstrated overall a high accuracy of the Annalise Enterprise CXR AI tool.

For a finding which is time sensitive and critical, such as pneumothorax, sensitivity must be emphasised over specificity, as the risks of missing the diagnosis of pneumothorax in general would outweigh the risks of overcalling pneumothorax (which would usually prompt further clinical review and possible CT investigation).

Sensitivity for rare findings is difficult to evaluate in datasets of general cases due to very low prevalence - for these cases, an emphasis has been placed on ensuring a sufficiently low rate of false positive (overcalls) AI results as it is well recognised that a high false positive rate (particularly when the false positives outnumber the true positive AI results) erodes clinician confidence in a model, and leads to erroneous dismissal of the true positive cases when they arise.

Analysis

Overcalling of pathology

There is general consensus of false positive tendencies for some pathologies such as rib fractures or peribronchial cuffing, which may be mitigated through general awareness and/or reducing sensitivity.

Several false positive AI findings of rib fractures are so close to the confidence threshold that decreasing the sensitivity threshold could lead to many missed detections. The current threshold is set at the minimum approved value, meaning that there is scope to increase threshold to decrease sensitivity. However, this carries the risk of missing further findings, in the setting of reports of AI false negative results for fractures. Rib fractures are usually straightforward to exclude on the images or can be correlated with no corresponding pain clinically. Further considering the low clinical significance of this false positive finding, current settings are kept for now. To determine with more confidence if it is appropriate to adjust thresholds, further analysis of a larger set of acute rib fracture findings to

identify how many were close to the threshold and then cross-checked against the radiology reports to confirm whether they were true or false positives can be conducted.

On the other hand, the threshold for peribronchial cuffing is already set at the maximum approved value, meaning that there is no scope to increase threshold without additional review and validation by the Clinical Research Team to consider clinical trade-offs between sensitivity and specificity (Greenhill, T., 2025).

Findings turned off to minimise confusion

Some AI findings have overlapping confidence intervals, resulting in conflicting AI findings. For example, “In position NGT” AI finding has been turned off to prevent confusion, as it is often listed concurrently with “Sub-optimal position”.

“Patient rotation” AI finding has also been turned off to minimise referral enquiries, as these are more frequently related to patient factors rather than technical issues.

No AI findings

13% of all CXRs analysed by Annalise have no AI findings; future audit of these to correlate with formal reports will yield useful data with regards to true negatives and assist in guiding the degree of confidence for a normal report in this setting.

Lung masses detection

One of the most impactful AI benefit is the detection of lung nodules, potentially resulting in early diagnosis and treatment of malignancies. Apart from improved patient care, the cost savings of avoiding more complex and prolonged treatment for delayed diagnosis of advanced pathology cannot be underestimated.

Survey

A post-trial survey will be conducted. This will also be compared with pre-trial survey. In the UK, an NHS survey 12 months into a similar pilot trial with secondary capture concluded that AI findings are extremely helpful second opinions for clinicians, improving efficiency and accuracy. Clinicians reviewing the AI findings on PACS prior to a formal report observed that they will often pick-up more subtle pathology that they may miss on low resolution emergency department computers, and that AI is helpful for preventing adverse outcomes from missed diagnoses particularly for less experienced staff.

Conclusion

According to the World Economic Forum (2022), much of the value in AI resides in choosing the right problems to solve. The key skill is the ability to understand the art of the possible with AI, while identifying the main risks it creates. With numerous AI tools emerging for radiology, selection of the most effective product to trial and implement is crucial in the setting of limited healthcare funds. Leaders should focus on the big picture and support change resistance in early adopters of AI, by conducting education and feedback sessions to engage stakeholders. It is important to do proof-of-concept pilots in advance of committing to the full roll-out of AI solutions to ensure upholding of AI ethical principles and value-based medicine. It is incumbent on all health professionals to be aware of the fast changing AI landscape, so that we can mitigate the risks and leverage on the benefits for patient care.

References

- Australian commission on safety and quality in health care, 2025, "AI safety scenario: Interpretation of medical images.", <https://www.safetyandquality.gov.au/sites/default/files/2025-08/ai-safety-scenario-interpretation-of-medical-images.pdf>
- Chen, C., & Cui, Z. (2025). Impact of AI-Assisted Diagnosis on American Patients' Trust in and Intention to Seek Help From Health Care Professionals: Randomized, Web-Based Survey Experiment. *Journal of medical Internet research*, 27(1), e66083-e66083.
- Department of Industry (Science and Resources), 2023, "Safe and responsible AI in Australia", <https://consult.industry.gov.au/supporting-responsible-ai>
- Greenhill, T., 2025, "Threshold analysis: peribronchial cuffing and acute rib fracture." Annalise Enterprise.
- Jones, C., 2025, "Royal Brisbane and Women's Hospital Annalise.ai Chest X-ray AI Project", RBWH.
- Mousa, D. (2025), "The algorithm will see you now", <https://worksinprogress.co/issue/the-algorithm-will-see-you-now/>
- Nong, P., & Ji, M. (2025). Expectations of healthcare AI and the role of trust: Understanding patient views on how AI will impact cost, access, and patient-provider relationships. *Journal of the American Medical Informatics Association : JAMIA*, 32(5), 795-799.
- Quinn, T. P., Senadeera, M., Jacobs, S., Coghlan, S., & Le, V. (2021). Trust and medical AI: the challenges we face and the expertise needed to overcome them. *Journal of the American Medical Informatics Association : JAMIA*, 28(4), 890-894.
- Soni, A., et al., 2025, "Survey on use of AI in Interpretation of CXR", Stockport, NHS Foundation Trust
- World Economic Forum, 2022, "Empowering AI Leadership", <https://www.weforum.org/publications/empowering-ai-leadership-ai-c-suite-toolkit/>.
- Yan, T. D., Jalal, S., & Harris, A. (2025). Value-Based Radiology in Canada: Reducing Low-Value Care and Improving System Efficiency. *Canadian Association of Radiologists journal*, 76(1), 61-67.

Appendix A: Excerpts from “Trust and medical AI: the challenges we face and the expertise needed to overcome them.” Quinn, et al. (2021)

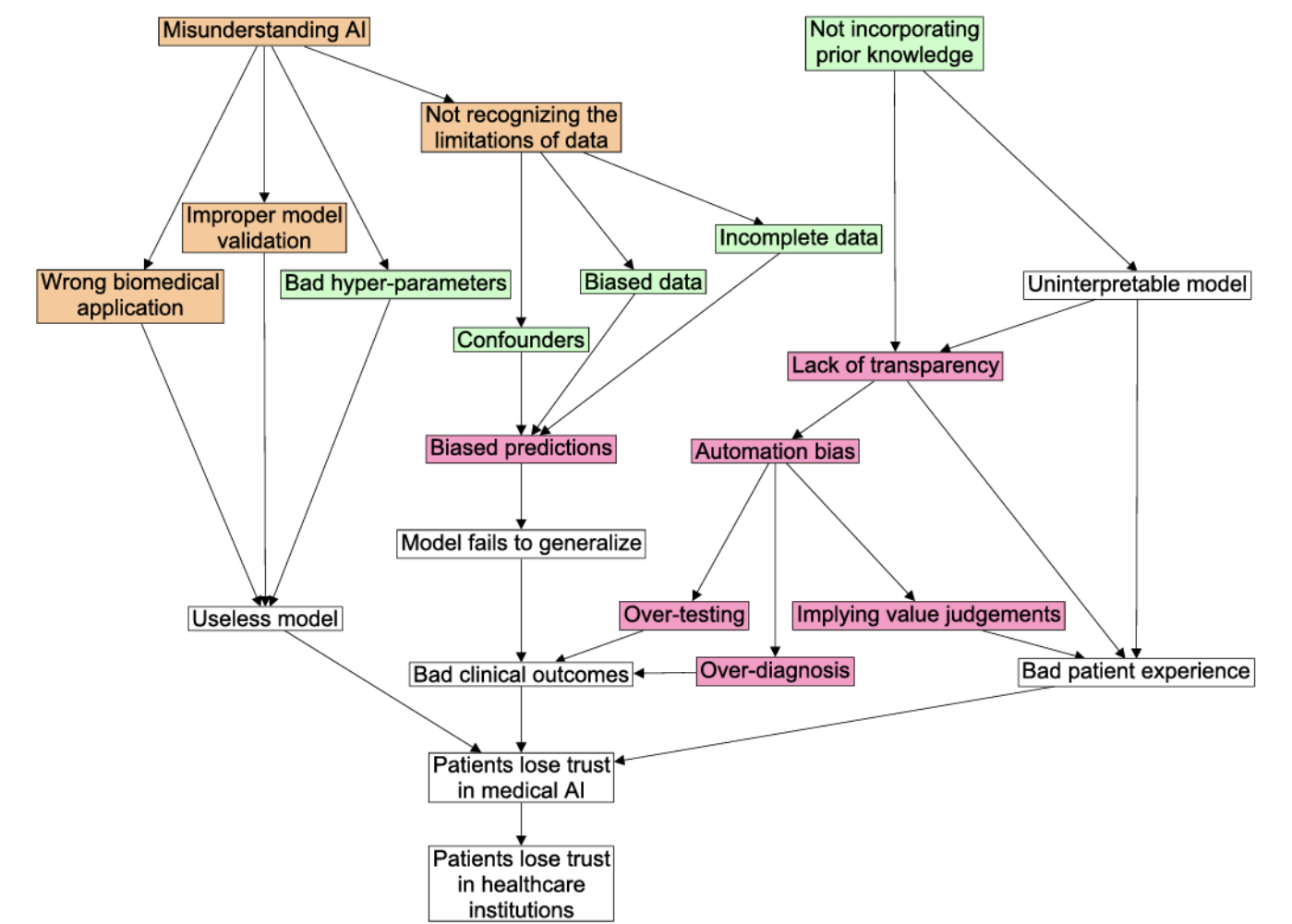
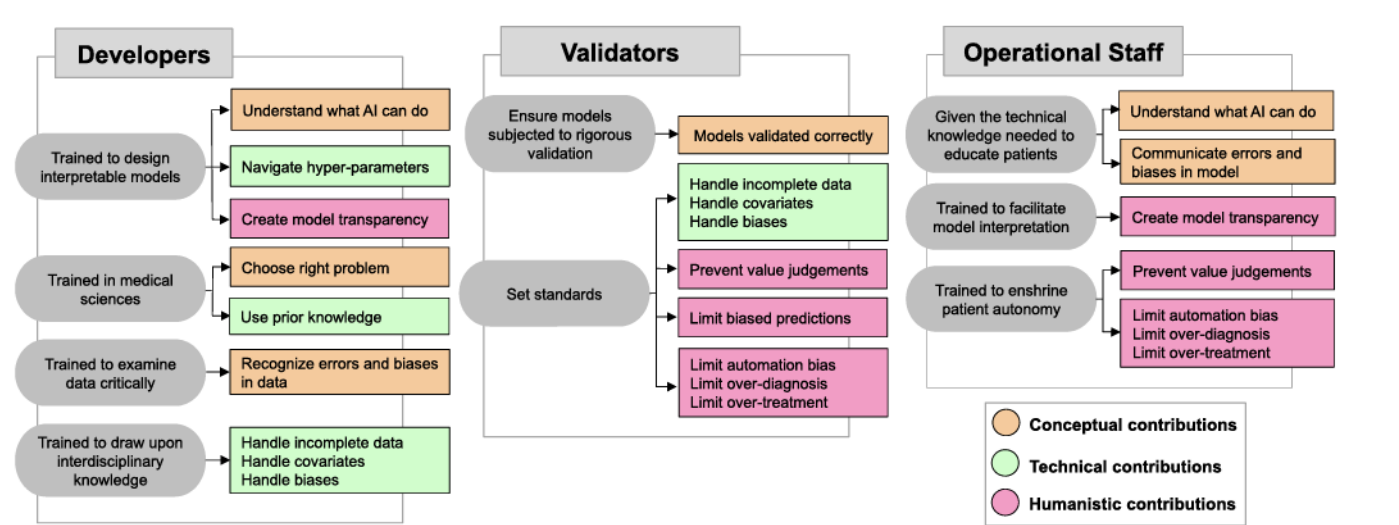


Figure 1. This figure shows how key challenges in medical AI relate to one another and to clinical care. If these challenges go unaddressed, the consequences could act concertedly to erode trust in medical AI, which could further undermine trust in our healthcare institutions. Node color represents the type of challenge: conceptual (orange), technical (green), or humanistic (pink). Uncolored nodes represent consequences.

Table 1. This table summarizes how accredited expert groups—developers, validators, and operational staff—can help overcome the key challenges in medical AI. Node color represents the type of challenge: conceptual (orange), technical (green), or humanistic (pink).



Appendix B: Acknowledgements

Annalise.AI pilot trial:

Dr Jason Brown – Chief Digital Health Officer RBWH (Project Executive & Chair)

Steve Gough – Principal Project Manager, Digital Metro North

RBWH Department of Medical Imaging - Medical and Informatics Team

Shared insights from South Australia Medical Imaging:

Associate Professor Marc Agzarian (SAMI Clinical Director) and John Hillier (SAMI Data Analyst)