AI Safety Scenario

Interpretation of medical images

This safety and quality scenario provides clinicians with an example application of Artificial Intelligence (AI) tools which use machine learning in medical imaging to demonstrate safe and responsible use.

The example is Artificial Intelligence (AI) supported detection of clinical findings in medical images, acknowledging AI is used in a variety of ways in the medical imaging sector. This guidance should be read in conjunction with the Commission's *AI Clinical Use Guide*. Also review the Australian Health Practitioner Regulation Agency (Ahpra) and the National Boards guidance and information from your professional association or college for example The Royal Australian and New Zealand College of Radiologists.

AI tools using machine learning (ML) have increasing use and can support clinical decision-making.[1] Examples include AI-based algorithms used in medical imaging systems to analyse mammograms to detect cancer and classify breast density and for abnormality detection in chest radiographs.

Although computer-aided medical image interpretation has been used for decades, there are key differences in contemporary AI tools:

- They are trained and tested to detect specific findings within defined populations and clinical settings.
- ML algorithms within AI tools may operate with limited transparency, making it difficult for clinicians to understand how they produced specific predictions or classifications.
- Are usually TGA approved as a medical device, including the ML algorithm.

The use of AI tools in medical imaging can have a direct impact on clinician decision-making and due diligence should be taken to consider appropriate use, consent, target population, known limitations such as bias, risks and benefits, ongoing monitoring and outcomes.

Before using AI for image interpretation

Review the evidence

As most AI tools used in medical image interpretation require approval by the TGA, the developer or your organisation should provide information about the AI tool's Australian Register of Therapeutic Goods (ARTG) inclusion status, intended use and instructions for use including risks and risk mitigations. Also carefully review and observe the identified 'intended user' as stated in the instructions for use. Considerations for assessing AI in image interpretation[1] include:

- **Relevance:** Confirm the problems that the AI tool is intended to solve, potential clinical or operational benefits, and clinical risks.
- **Performance and validation:** Confirm the tool is trained and tested on data aligned to the target patient population, target anatomy, imaging equipment and data quality requirements. AI algorithms are trained and tested in specific conditions and with particular data and rare conditions may present challenges for AI tools due to limited training, validation and testing data.
- **Usability and integration:** Understand how the AI tool will be used in your workflow and where you as the responsible clinician provides final authorisation of the result (initial or secondary viewer application each carry risks). Understand and mitigate automation bias and how the AI tool is integrated with other systems and your workflow.
- **Monitoring:** Know how to identify poor performance, such as false positives and false negatives and what to do when they occur.

Establish AI transparency and informed consent procedures

Clinicians are responsible for the safe and appropriate use of AI tools that support the interpretation of medical images. They should be aware of how the tool impacts patient management.

Clinicians should be prepared to:

- Explain the benefit of improved abnormality detection
- Review risks to the patient including AI bias associated with training data suitability to the patient
- Disclose the use of AI in line with your organisation's approach. For example, posters in waiting rooms, information on registration forms or during the appointment booking process (on-line or in person).

AI ML image interpretation tools generate data that is regarded as personal and sensitive information under the Privacy Act 1988 (Cth) and pragmatically, must be stored and processed in Australia (see *AI Clinical Use Guide*).

Requirements for consent are outlined in the *AI Clinical Use Guide* including where AI outputs can influence clinical decisions and privacy risk. Consent should encompass a risk-based approach and incorporate both clinical and non-clinical benefits and harms including privacy, data handling and use and where applicable sharing with third parties. The nature of consent will be situational and largely determined by your organisation.

- Understand how consent will be documented for example, on registration forms (paper or electronic), as part of online booking terms and conditions, through privacy collection notices with opt-out options, in the patient's clinical notes.
- Determine how circumstances where consent is not given will be managed. The approach should be common across all clinicians using the tool.

Understand and manage common limitations and risks

All AI image interpretation outputs must be reviewed for accuracy and quality (refer to the *AI Clinical Use Guide* and ACSQHC standards).

Clinicians are susceptible to automation bias when they accept AI image interpretation as being accurate and complete without adequate review.[2]

AI tools should be updated through managed processes. Updates can change analysis, interpretation of images and final outputs. Therefore, it is critical to establish or be aware of the governance mechanisms in place within your organisation to assess the ongoing performance of the AI tool and how patient information is used.

- Confirm with you organisation what governance and management structures are or will be put in place to assess and improve the performance of the AI image analysis tool.

When using AI for image interpretation

Continue being transparent with patients

- You should advise patients how the AI tool supports care delivery, where in the care process the AI will be used and where the clinician is involved in determining results.
- Provide information about the risks and benefits of using AI tools.
- Outline the risks and limitations of AI tools, how these issues are managed, and provide an opportunity for the patient to discuss the use of AI and their alternatives.

Be satisfied with consent

- You should be satisfied that appropriate and proportionate consent procedures have been followed and appropriately documented in line with your organisation's procedures.
- Where applicable, ensure consent covers the disclosure of personal and sensitive information to third parties and the purpose of sharing, including where personal and sensitive information is planned to be used for training AI tools.

Review AI image interpretation

Responsibility for detection, diagnosing and reporting remain with the clinician and all AI outputs must be checked for accuracy and quality.

- Always ensure you are using the AI tool as intended (target patient, target anatomy, etc.).
- Always review AI outputs for accuracy and quality and make any updates or corrections for final reports/diagnosis as necessary.
- Final results and reports must meet medico-legal requirements with the understanding that reports may be shared with other organisations or clinicians.
- Label records/results indicating AI was involved in its creation when appropriate.

After using AI for image interpretation

Monitor and evaluate performance for safety and quality

Undertake regular reviews of how the AI image analysis tool is performing and its impacts on the quality and safety of clinical care. Refer to the *AI Clinical Use Guide* and '*Understand and manage common limitations and risks*' section.

- Discuss with your colleagues and in clinical governance and/or management forums common observations both positive and negative.
- Where performance is deteriorating discuss with your employer or AI developer to determine how the AI training can improve performance.
- If there is an issue with the results of a medical imaging system including AI issues, that have caused an adverse event, follow your organisation's process to report the issue to the TGA.
- Report perceived breached of medical device regulations, or questionable practices to the TGA and report incidents and near misses you observe to the TGA directly as per your organisation's reporting requirements.

References

1. Magrabi F, Bates L, Brooke-Cowden K, et al. 2024, Literature review and environmental scan report: AI implementation in hospitals: legislation, policy, guidelines and principles, and evidence about quality and safety.
2. Coiera E. (2015). Guide to health informatics. Taylor & Francis.

For more information

Please visit: safetyandquality.gov.au