Australian Commission on Safety and Quality in Health Care

Australian Government
Department of Health, Disability and Ageing

AI Clinical Use Guide

Guidance for clinicians

Artificial Intelligence (AI) tools can support a wide range of clinical tasks, improving care delivery, health outcomes, and patient satisfaction.[1]

Introduction

AI has supported healthcare in Australia for decades, including early clinical decision support systems used to identify medicine interactions. Like all clinical innovations, AI can bring significant benefits for patient care but also introduces new risks.[2] The rapid advancement and adoption of AI can result in new and increased risk, especially as evidence of safety and efficacy may lag behind implementation.[2] This guidance and associated clinical scenarios support clinicians, together with their patients, in using AI safely and responsibly in patient care and are structured to support the steps of 'before you use', 'while you use' and 'after you use' AI tools.

As with all healthcare technologies, clinicians must meet their professional and legal obligations, including Australian Health Professionals Regulatory Authority (Ahpra) and National Boards guidance in relation to patient safety and best practice in the application of AI tools. This requires you to:

- Understand how the AI tool will be used in your workflow and recognise your accountability for all AI outputs that inform a clinical decision, finding or documented record.
- Understand the problems that the AI tool is intended to solve, potential clinical or operational benefits, and clinical risks.
- Confirm the evidence base for the tool including its risks and limitations and how these are managed.
- Be prepared to discuss the potential benefits, risks and limitations of an AI tool with patients.

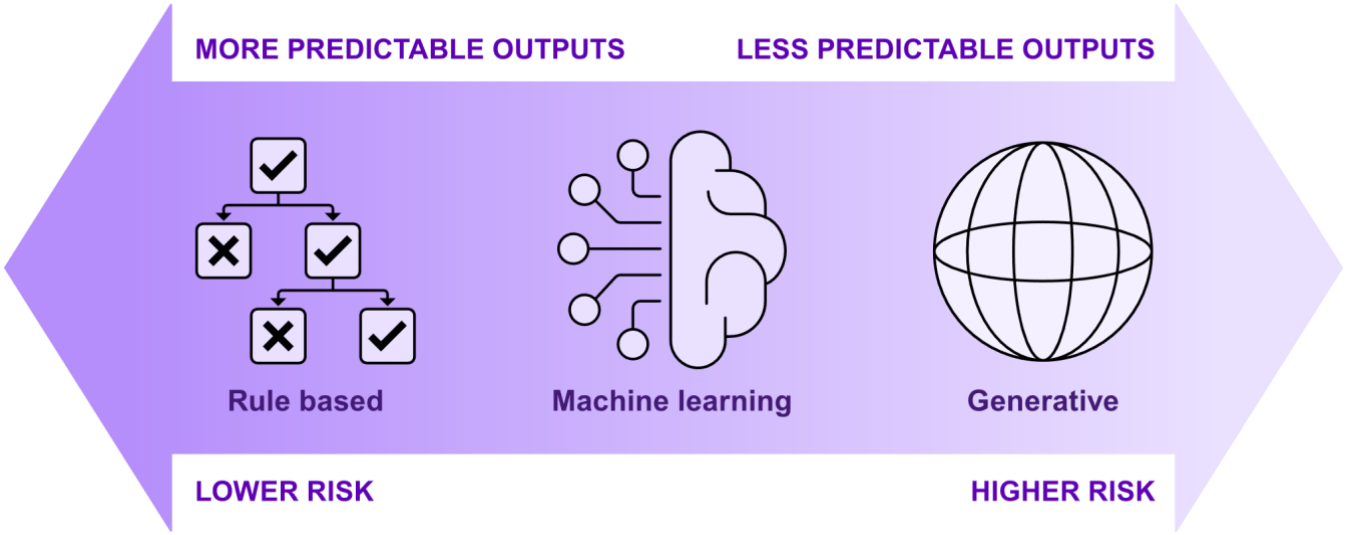
- Recognise ethical implications including the potential for inequity and bias in how AI tools use and process data and generate outputs such as recommendations.
- Know that an AI tool meets the definition of a medical device when it is used for diagnosis, prevention, monitoring, prediction, prognosis, treatment, or alleviation of disease, injury or disability and must therefore comply with Therapeutic Goods Administration (TGA) regulations.
- Educate yourself on how AI tools operate, either through your organisation or through external avenues.
- Comply with the governance and oversight for AI in your organisation. Smaller organisations may need to establish appropriate governance and processes to ensure AI tools are deployed safely and responsibly and in line with their intended use.

Also review information that may be available from your professional association or college.

Before you use AI tools

While AI tools are often implemented organisation wide, it is essential for each clinician to understand how the technology might impact care delivery and patient outcomes. AI tools in healthcare may incorporate different computational methods (**Figure 1**). These include simple rules to analyse clinical information and provide recommendations (for example, identifying patient risk factors or medication interactions); machine learning models to assist with diagnosis (for example, in medical imaging); and generative AI to summarise information (for example, ambient AI scribes for creating clinical notes). It is important for clinicians to understand that new AI technologies incorporating machine learning and generative AI pose substantially different risks to rules-based products that have been in the market for much longer.

Figure 1: AI tools may incorporate one or more computational methods



AI tools are increasingly combining multiple computational methods and technologies. To safely integrate AI into clinical workflows, clinicians must familiarise themselves sufficiently with the intended use of each AI tool, to understand the benefits and potential harms.

Critically assess scope of use and the available evidence

When a new AI tool is introduced to clinical practice, its limitations and risks must be critically assessed for safety and efficacy impacts. AI development often occurs in highly controlled and idealised settings which can outpace the development of robust evidence in 'real-world' clinical settings. It is essential to review available evidence, such as published literature, required medical device labelling information (including instructions for use) and information from the AI developer supporting intended use, accuracy, efficacy, and safety. In cases where supporting evidence for the performance of AI tools is limited or absent, including non-TGA regulated AI tools, clinicians together with their patient must carefully and transparently weigh the potential harms against the anticipated benefits prior to use in clinical settings. Refer to the Commission's Introduction of new interventional procedures and clinical practice innovations for further guidance.

If you are not satisfied with the evidence to support the accuracy, efficacy and safety of the AI tool, you should re-consider its use in consultation with your organisation.

Establish AI transparency and informed consent procedures

Clinicians are responsible for ensuring that patients are informed about the use of AI in their care, including its purpose, scope of use, and the associated benefits and risks. Be prepared to explain:

- That you review outputs and remain responsible for final clinical documentation and decisions
- Why the AI tool is being used in the delivery of care
- Its benefits, limitations and potential risks
- The processes in place for monitoring safety and performance of AI tools.

The broad and varied uses of AI make consent complex and there is no single approach. The method of disclosure and type of consent required will normally be determined by your organisation and depend on the nature of the AI tool, the clinical and non-clinical concerns or issues including privacy, data handling and use, and where applicable, sharing with third parties. You should be:

- Satisfied that appropriate and proportionate informed consent procedures and documentation are in place and adhered to; and contribute to continuous improvement processes
- Adopting your organisation's approach for managing situations where a patient does not consent to the use of AI in their care.

Understand common limitations and risks

Machine learning and generative AI tools are shaped by the data they are trained on, known as training data, and by their nature create information that may not be factual, sometimes called 'hallucinations'.[3] If underlying AI tool data and computational models lack diversity or is not representative of the patient cohort, it can lead to biased outputs. Over time, these outputs may also change. There are documented cases where AI tools have disadvantaged certain patient groups due to underrepresentation in the training data. This bias raises important ethical and equity concerns, along with potential clinical risks such as inappropriate treatment recommendations, inaccurate healthcare records and diagnostic errors. You should be aware of the following when using AI tools:

- **Performance:** Evidence about the performance of AI tools in clinical settings may lag behind their implementation. Establish or contribute to governance forums to monitor AI performance within your organisation.
- **Data representation:** Accurate outputs depend on AI tools being trained on data that reflects the diversity of the patient population. Ask the developer about the dataset used to train the AI tool and ensure it can be used on the local population.
- **Fairness and equity in AI training and outputs:** AI tools should be inclusive and accessible and their use should not involve or result in unfair discrimination against individuals, communities or groups. Biases in AI outputs can arise from ethnicity, sexuality, disability, and age represented in training data. Individual patient characteristics such as co-morbidities may not align with the data the AI was trained on. Treat the AI as a decision support tool. If AI outputs contradict clinical intuition, investigate further.
- **Patient data use and consent:** Some AI tools may use patient information for ongoing training and share it with third parties. This could risk future re-identification or a data breach. Data sharing and/or potential secondary use require explicit consent and must also be disclosed to patients during consent processes. Confirm your organisation's consent procedures and documentation (see ACSQHC – Informed consent).
- **Medical devices:** AI tools listed on the Australian Register of Therapeutic Goods (ARTG) will have instructions for use including known limitations, risks and mitigation steps. Review instructions before use.

Meet privacy requirements

When using AI, clinicians need to meet the requirements of the Privacy Act 1988 (Cth) for the storage and processing of personal health information. You should:

- Confirm with your organisation or AI developer that personal and sensitive information is stored and processed in Australia. Australian privacy law allows personal information to be processed and stored overseas only when equivalent privacy and security standards are met (Australian Privacy Principle 8) however, it is mostly impractical for a clinician to determine equivalency.
    - If personal or sensitive information is used for ongoing AI training, confirm with your organisation how informed patient consent is obtained and documented.

Ensure ongoing support for safe and secure use

AI tools, like other digital health technologies, require ongoing support and monitoring to ensure safe operations and security of information. You should:

- Only use AI tools that are authorised and/or supplied by your organisation to ensure standard technical support, maintenance, supporting infrastructure/hardware and security management including cyber-security approaches are in place. This applies to clinicians that have both clinical and managerial responsibilities in sole-trader or smaller organisations. Note that generic, publicly available AI tools may not have been designed and trained specifically for health care purposes.
- Confirm or establish the forums where AI tool performance can be measured and discussed as AI tool performance can change over time. Examples are existing clinical, incident management or technical governance forums, or practice level management meetings.

- Escalate issues that have serious patient consequences stemming from AI use. Channels for escalation include:
    - Reporting adverse events for medical devices to the TGA
    - The relevant privacy regulator in your state or territory, or the Office of the Australian Information Commissioner (OAIC) for breaches of personal and sensitive information
    - The relevant healthcare complaints authority in the case of actual or potential harm.

While using AI tools

Use AI safely and responsibly

- You must use AI safely and responsibly, applying professional judgement and always exercising due care in the context of your patient. You remain responsible for confirming the accuracy and completeness of records, results or documentation that will be acted on by you, other clinicians or patients, now and into the future.
- If the AI tool is TGA approved, read and understand the instructions for use.
- Always review AI generated outputs for accuracy and quality in a timely manner, ideally during or at the completion of the patient interaction and in the context of the patient:
    - Correct any errors in summaries, reports or analysis and save clinician authorised information in the relevant healthcare record and/or analysis system.
    - Review alerts and notifications for action, then record your action. For example:
        - Generative AI tools have been found to create fictitious outputs and/or omit important clinical information or 'mishear', for example in the summarisation of patient consultations (see RACGP information).
        - Machine learning AI for image interpretation may produce false positive or false negative interpretations.
        - Computer generated alerts and notifications may not account for the specific circumstances of your patient.

Understand the risk of automation bias

Research shows that ongoing use of digital technologies is associated with *automation bias* which in the case of AI is the tendency to over-rely on AI tools by prioritising the output of the AI tool over clinician's own independent judgement. Automation bias can lead to two types of errors: acting on incorrect recommendations (errors of commission) and failing to act when AI misses something (errors of omission).[4] This can result in incorrect or missed diagnoses, inappropriate treatments, overlooking important patient information. You must always critically evaluate AI outputs, recognising that these tools support but do not replace clinical judgment.

Facilitate transparency and disclosure AI use

- You should advise patients how the AI tool supports care delivery and where in the care process the AI will be used.
- 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 the patient the opportunity to discuss the use of AI and alternatives.

Be satisfied with consent

- You and your patient should be satisfied that appropriate and proportionate consent procedures relative to risks and benefits have been followed and appropriately documented in line with your organisation's procedures.
- Where appropriate, explain 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 and ensure an environment where the conditions for informed consent are met (decision-making capacity, voluntary and provision of information).
- Where applicable, ensure consent covers the disclosure of personal information to third parties and the purpose of sharing that information including where personal and sensitive information is planned to be used for training AI tools (see *'Meet privacy requirements'* section).

Maintain patient privacy

As with clinician generated healthcare records, AI tools are likely to gather, use, or generate personal and sensitive information. This information should be treated in the same way as other sensitive health information that may be sent to other clinicians, organisations or health record systems. You should ensure that your work practices comply with all relevant policies, codes, and laws (for example, restrict use of technologies to approved medical devices; only transmit AI data via approved secure channels and take reasonable steps to destroy or de-identify the personal information held once the personal information is no longer needed).

After using AI tools

Monitor and evaluate performance for safety and quality

Continuous monitoring is required by clinicians for the long-term safety and efficacy of AI technologies. You should:

- Ensure that records or results created with AI tools meet the same quality standards as those created manually (see ACSQHC documentation standards).
- Label records indicating that AI was involved in their creation.
- Escalate any identified risk, near miss or actual harm to patients immediately within your organisation and contact the patient to assess immediate next steps to reduce the risk or impact of harm. Escalation outside your organisation to a regulatory body may be required in certain circumstances (for example, the OIAC's notifiable data breach). See '*Ensure ongoing support for safe and secure use*' section.
- Ensure all AI outputs are continually monitored over time for accuracy and relevance, and quality improvement.
- Be aware that updates to AI tools may not be obvious and outputs may affect downstream functions and systems.
- Be alert to changes in the functional scope of AI tools which can change or 'creep' over time through software updates or upgrades. You may not be notified when these changes occur. Such changes are likely to impact the intended use and evidence supporting its use and may result in the tool being classified as a medical device requiring listing on the ARTG. For example, if an ambient AI scribe adds recommended tests or treatments to a summary of a consultation it is a medical device subject to regulation. See '*Critically assess scope of use and the available evidence*' section.

- Report perceived breaches 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.

- **Support AI tool quality assurance and monitoring by contributing to organisational level clinical governance forums on the ongoing performance of the AI tool and where appropriate to the AI developer.**

References

1. Productivity Commission 2024, Leveraging digital technology in healthcare, Research paper, Canberra.
2. 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.
3. Coiera E, Fraile-Navarro D. AI as an Ecosystem – Ensuring Generative AI Is Safe and Effective. NEJM AI. 2024;1(9):AIp2400611.
4. Coiera E. (2015). Guide to health informatics. Taylor & Francis.

For more information

Please visit: safetyandquality.gov.au