

AI in Guidelines: An Australian Lens

Report from the workshop held as part of [Living Evidence 2026 Symposium](#)

Executive Summary

The *AI in Guidelines: An Australian Lens* workshop, held as part of the [Living Evidence 2026 Symposium](#), brought together approximately 70 Australian guideline developers, researchers and partner organisations to explore how artificial intelligence (AI) can support guideline development. Guideline developers recognised that AI is already being used across parts of the guideline development process particularly for high-volume, repetitive and administrative tasks.

Participants emphasised the importance of ensuring AI should not be used as a replacement for expert judgement, but rather to assist teams in undertaking structured, rules-based tasks (for example, meeting minutes, de-duplication, summarising, screening support, and evidence surveillance). The greatest advantage to guideline development would be reducing workload pressures and creating capacity for additional activities that require deliberation, judgement, or clinical and methodological expertise. Resources could additionally be redirected to other value-add activities that improve engagement with interest holders, tailor dissemination, support implementation and evaluate impact.

Participants identified several risks with the use of AI that could undermine trust and credibility if unmanaged, particularly hallucination, inaccuracies, inconsistent outputs, and privacy or confidentiality concerns. Participants noted that some of these risks could be managed by using a “**human-in-the-loop**” approach, always treating AI outputs as draft material and ensuring outputs are subject to full review, verification and sign-off processes. Further reflections from workshop participants emphasised:

- **using AI strategically** by prioritising its use for tasks that are high burden (for example, automation of repetitive work) without compromising methodological quality.
- **embedding safeguards for responsible use** by protecting confidentiality (for example, closed systems where required, de-identification and clear disclaimers), actively managing bias, and maintaining credibility through oversight and checks for known AI limitations.
- **strengthening governance and transparency** by developing standards and disclosure requirements to clarify acceptable tools, data handling expectations, validation requirements, and how AI assistance is reported in outputs.
- **building capability** by improving AI literacy and workforce confidence through training, shared learning opportunities and practical guidance to enable safe, effective and equitable uptake.

Overall, workshop participants were cautiously optimistic about the use of AI in guideline development but wanted further support in understanding the risks. Strong governance frameworks, standards and transparent disclosure of AI use were viewed as essential to building confidence in users, with a focus on ensuring the source data is understood and its application used appropriately and consistently across activities.

Summary of the day

The *AI in Guidelines: An Australian Lens* workshop was held on 18 February 2026 as part of the three-day Living Evidence 2026 symposium hosted by Cochrane Australia.

The workshop was hosted by the Guideline International Network Australian and New Zealand Regional Group (GIN-ANZ) and supported by JBI, the National Health and Medical Research Council (NHMRC) and Cochrane Australia.

The workshop brought together national and international experts in guidelines and evidence synthesis to discuss and collaborate on methods and approaches to integrate AI into evidence synthesis and guideline development.

It was attended online and in person by approximately 70 participants from across the guideline development community in Australia, including guideline developers and researchers from various universities and institutions.

Objectives

The key objectives of the workshop were to:

- identify how AI is currently being used to support guideline development processes
- identify opportunities, challenges, and requirements for ensuring the safe and responsible use of AI in Australian guideline development
- discuss the future role of AI in developing Australian guidelines for clinical practice.

Presentations

The day commenced with a series of presentations:

- **AI in guidelines: considering equity, fairness, and bias** – Professor Steve Wesselingh, NHMRC, Australia
- **AI in Guidelines from Concepts to Practice** – Assistant Professor Bernardo Sousa Pinto and Assistant Professor Manuel Cruz, Faculty of Medicine of the University of Porto, Portugal
- **Benefits of AI in guidelines: Let's trust the machines?** – Adjunct Associate Professor Britta Tendal Jeppesen, Future Evidence Foundation, Australia
- **Challenges with AI in guidelines: Should we trust the machines?** – Professor Zachary Munn, Adelaide University, Australia
- **How AI-enabled guideline developers can leverage Evidence Synthesis Infrastructure Collaborative (ESIC)** – Professor John Lavis, ESIC
- **Collaborating for Responsible Use of AI: From Evidence Synthesis to Guideline Development** – Dr Kylie Porritt, JBI, Adelaide University, Australia
- **Building AI Enhanced Workflows for Living Guidelines for Pregnancy and Postnatal care** – Dr Shaira Baptista, Monash University, Australia

Panel discussion

Following the presentations, a panel discussion facilitated by Professor Tari Turner - Director of the Australian Living Evidence Collaboration, Monash University - asked the following panellists about priorities for AI use in Australian guideline development:

- Professor Steve Wesselingh, NHMRC, Australia
- Adjunct Associate Professor Britta Tendal Jeppesen, Future Evidence Foundation, Australia
- Professor Zachary Munn, Adelaide University, Australia
- Dr Kylie Porritt, JBI, Adelaide University, Australia
- Dr Shaira Baptista, Monash University, Australia

The discussion explored several critical topics, including:

- the current readiness of AI tools for use in guideline development
- how to ensure guidelines are trustworthy and maintain high-quality standards with the use of AI - particularly balancing efficiency gains with methodological rigour
- who holds responsibilities of managing the risk of AI outputs - particularly considering the end goal of guidelines informing clinical practice where there are important consequences for accountability
- the extent of human oversight and expertise in ensuring AI is used appropriately and responsibly
- how to ensure AI does not widen inequities
- the value in establishing a community of practice for guideline developers or further opportunities for collaboration and shared infrastructure to support AI integration across the Australian guideline community.

The panel also highlighted the need for clear governance, standards, and guidance to support the safe and ethical use of AI in guideline development in the Australian context. Several important questions for future consideration were raised:

- how do we ensure we don't widen inequities,
- who holds responsibilities of managing the risk of AI outputs - particularly in clinical practice,
- how to ensure guidelines are trustworthy with the use of AI.

Workshop activities

Attendees participated in a series of workshop activities to focus on practical considerations of AI in guideline development.

Activity 1: SWOT analysis

Attendees used a strengths, weaknesses, opportunities and threats (SWOT) analysis framework¹ to map the current use of AI in the development of guidelines. As part of the SWOT analysis, attendees also explored advantages, limitations and risks associated with using AI in guideline development.

Activity 2: Small group discussion – components of guideline development

The second session of workshop activities explored AI capability across the components of the guideline development process. This was explored through a series of rotating small discussion groups focused on the various components of guideline development:

- Scoping and prioritisation
- Evidence surveillance and living guideline maintenance
- Recommendation development
- Public consultation
- Governance

Activity 3: Small group discussion – shaping the future of AI in guideline development

Finally, participants considered the future of AI in guideline development through small group discussions. Discussion topics included implementation, capacity building and future priority areas.

Key themes for each activity are summarised in Table 1 below.

¹ A SWOT analysis is a strategic planning tool used to identifying **Strengths** and **Weaknesses**, as well as external **Opportunities** and **Threats**

Outcomes from the workshop activities

1. SWOT analysis: the current status of AI use in guideline development (an Australian lens)

Participants identified current approaches, opportunities and threats (SWOT) with the use of AI in guideline development in Australia through a SWOT analysis outlined in **Table 1**.

Table 1: Outcomes of SWOT analysis: AI in current guideline development (an Australian lens)

Strengths	Weaknesses	Opportunities	Threats
AI reduces time spent on repetitive or lengthy administrative tasks (e.g. minutes, public consultation submission management, project management, compliance to style guides). AI is able to extract useful insights from large volumes of information (e.g. scoping, collating existing guidance, evidence synthesis). AI can free up resources for more complex tasks that require more human judgement. AI can provide a starting point for work and supports idea generation (e.g. drafting initial guideline content). AI-enabled evidence maps may allow for better policy decision making. AI performs well with language translation (e.g. multiple languages, or translation into plain language advice from complex clinical advice).	Quality of AI outputs are uncertain with instances of hallucination, inaccuracies and poor reproducibility. Time and effort are often still required to review and edit AI content. Lack of confidence and low AI literacy are barriers to adopting AI more widely. Overcoming this requires training and plans to integrate AI into workflows (e.g. training on prompt engineering). Workforce capability poses a barrier to effective use of AI. This requires time commitments, dedicated resources and ongoing investment in upskilling the workforce.	Automation and speeding up of repetitive, manual, or lengthy tasks that involve large data sets, screening and identifying discrepancies. Living reviews could be more feasible with AI support to monitor new evidence and enable quicker updates. Identifying gaps in research, policy and equity to improve quality and relevance of guidelines and derivative products. Specialised tools and prediction models suited to guideline development and the health industry to improve consistency. Improved accessibility and communication through generating summaries, converting guideline style and format to suit different audiences. Development of derivative implementation tools (e.g. consumer resources including shared decision-making resources or conversation scripts).	Reduced quality of analysis (e.g. hallucinations and non-reproducible outputs). Risks to trust and credibility if AI outputs are not transparent, validated or well understood. Clear governance and human verification are needed. Privacy and confidentiality threats due to potential mismanagement of sensitive data in AI systems. Reduced critical thinking and intellectual capability from over-reliance on AI. Planetary health and environmental impacts have yet to be fully understood. Reduced funding and unrealistic timelines for guidelines as funders may expect work to be completed rapidly and at lower costs despite the need for verification of accuracy of AI outputs. Risks to job security and workforce pathways, particularly lower-skill roles, and uncertainty of how to best educate and upskill the workforce (to be able to critically assess AI outputs) where AI replaces these lower skilled pathways.

2. AI capability across the components of the guideline development process

AI tools currently being used to support guideline development in Australia

Throughout the workshop, participants discussed different AI tools they have used to support guideline development. A list of AI tools currently in use for guideline development in Australia, alongside details on the type of AI, and which guideline development stage(s) they may be useful for is available in [Appendix 1](#). Although the tools listed are not endorsed products, they illustrate the variety and possibilities that AI offers to support guideline development.

3. Key findings by theme

Participants were asked to focus on three topics: implementation, capacity building and future priority areas. From this discussion participants discussed several issues that should be considered with the application of AI and offered suggestions for its use in guideline development.

Key themes from the AI in guidelines workshop discussions

Implementation

Considerations	Practical application
- AI is most valuable when it reduces workload associated with manual high-volume, repetitive or administrative activities. - Adopting AI could allow for teams to spend more time on complex tasks that rely on nuanced human judgement such as panel deliberation, recommendation writing, dissemination/implementation efforts and impact evaluation. - Acknowledge that outputs can be inconsistent (for example, a user may input the same prompt into two different AI tools and receive different results), particularly if the model makes up information, or refers to out-of-date information. Outputs still need human review and verification.	- AI may be considered where it supports and streamlines components of the guideline development process (e.g. administrative and structured tasks such as drafting meeting minutes, managing documents, formatting outputs, and supporting project coordination; and components of the evidence synthesis process, such as searching, and assisting with screening activities). - When using AI to always assume there will be errors, and to treat AI outputs as draft material or starting points rather than final products. - Apply appropriate validation, verification, and expert oversight to ensure accuracy, transparency, and reliability - a human-in-the-loop approach is essential. - Set parameters around where the AI tool or model gets its information from and determine what information sources are eligible for inclusion - Undertake test runs and re-validate information when models, prompts or protocols change.

Capacity building

Considerations	Practical application
- Many potential users were cautiously optimistic about the use of AI in guideline development but noted a lack of AI literacy generally and confidence in using AI. - Participants emphasised the need for foundational skills, training, awareness of tools and support to keep up to date with changes.	- Consider shared learning opportunities through a community of practice - Standardise good practice guidance or standard operating procedures that is regularly updated - Endorse tool repositories and tools certified by authority organisations

Future priorities

Considerations	Practical application
Participants had reservations about the reliability of AI and wanted further information on the risks of using it (confidentiality, privacy and data security, copyright violations and bias in AI software owned by private companies) including how to minimise risks.	- Explore what rules, regulations and cultural practices (which could differ across institutions) could facilitate or hinder adoption of more widespread AI use. - Develop governance frameworks, policies on AI use, transparency and alignment with regulation and local standards. - Provide specific information for guideline development, particularly what tools were acceptable for use, what data can be used, where outputs are stored and how AI should be disclosed in final products

Conclusion

The *AI in Guidelines: An Australian Lens* workshop highlighted both the significant potential and the important limitations of AI in guideline development. Participants recognised that AI is already contributing to selected components of the process, particularly those that are structured, repetitive, or administrative in nature. When used appropriately, AI appears to offer opportunities to improve efficiency and reduce the burden on guideline development teams.

At the same time, AI is not a substitute for expert judgement. Concerns regarding accuracy, reliability, bias, and transparency, particularly within the evidence synthesis component of guideline development, highlight the need for a cautious and risk-managed approach. Maintaining a human-in-the-loop is essential to ensure that all AI-supported outputs are subject to rigorous validation, verification, and expert oversight.

Realising the benefits of AI in guideline development will require continued investment in governance, capability building, and shared standards. This includes developing clear guidance on the appropriate use of AI in guidelines, strengthening workforce confidence and literacy, and promoting transparent practices. Safeguarding data privacy and maintaining methodological integrity will be critical to sustaining trust in guideline processes and outputs.

Appendix 1 – AI tools

AI tools used in guideline development

Most major stages of guideline development

Tool name	Type	Description
Microsoft 365 Copilot	LLM	Microsoft's AI assistant integrated across the Microsoft Office Suite. - Generates text and image content, automates tasks - Can act as an administrative assistant.
ChatGPT	LLM	ChatGPT is a conversational AI chatbot created by OpenAI. Can understand questions, hold conversations, generate human-like text and image content, ideas, offers translation, organising files and chat logs.
Claude	LLM	Claude is an AI assistant (conversational chatbot) created by Anthropic, designed to help with complex thinking, writing, coding, and analysis. Generates writing, coding, research, math, analysis; strong at managing long inputs.
Gemini	LLM	Gemini (AI) is Google's family of advanced artificial intelligence models. - Powers Google's AI assistant. - Designed for reasoning, multimodal understanding, coding, and creative generation. - Generates text, images, audio, video content and code.
Perplexity	LLM	Perplexity is an AI-powered "answer engine". - Retrieval-augmented LLM tailored for research; helps with analysis. - Combines real-time web search with large language models to give direct, cited answers instead of long lists of links. - Can use focus modes to filter sources (e.g., academic-only).
Zoom AI Companion	LLM	Zoom AI is a companion that provides real-time assistance across the Zoom ecosystem. - AI meeting assistant. - Note taking, summaries, transcription, real-time question answering.

Tool name	Type	Description
Vibe AI	LLM	Vibe AI is a contextual AI workspace that links inputs into a continuously updated ‘memory’. • Captures information from meetings, conversations, documents and connected tools, transforming inputs into clear insights and actions. • Real-time transcription, time tracking, meeting notes, summaries, action items.
Miro	LLM ML NLP	Miro is an online visual collaboration platform built around an infinite whiteboard. Miro uses AI features to generate and transform board content. • AI-assisted content generation through creating sticky notes, mind maps, diagrams or draft documents. • Organises board content by scanning sticky notes and collating information and themes to generate insight summaries.

Stages of guideline development involving evidence review (e.g., scoping review, evidence review)

Tool name	Type	Description
Open Evidence	LLM	Open Evidence is an evidence-based clinical AI platform. • Retrieval-augmented AI; restricted to licensed, peer-reviewed clinical literature. • Tailored for clinicians, provides medical information, responds to medical questions and includes direct citations to evidence.
Consensus	LLM	Consensus AI is an AI-powered academic search engine. • Retrieval-augmented AI; tailored for academic research. • Searches and analyses peer-reviewed literature; provides summaries and citations.
TERA	ML NLP	TERA is a suite of automation tools designed to help researchers conduct evidence reviews and evidence synthesis more quickly. • Assists with protocol and methods writing, search strategy and syntax translation, de-duplication, screening, citations and analysis.
Otto SR	LLM	Otto SR is an AI-powered platform that automates key stages of systematic reviews. • Assists with search, screening, full-text retrieval and data extraction.

Tool name	Type	Description
EPPI-Reviewer	ML NLP	EPPI-Reviewer is a web-based evidence synthesis platform with built-in machine learning tools. - Trained on labelled datasets and works to improve efficiency of screening and review workflows. - Automates deduplication, study clustering, screening, study classification and monitoring of new literature for updates.
Covidence	ML NLP	Covidence is a web-based platform with built-in machine learning features that streamline systematic reviews. Filters studies, imports references, screens titles, abstracts and full text, data extraction, populates risk of bias tables.
Rayyan	ML NLP	Rayyan is an AI-assisted systematic review platform that uses machine learning to screen and prioritise studies. Organises and selects studies for systematic and scoping reviews.
DistillerSR	ML NLP	DistillerSR is an AI-enabled literature review and evidence management platform. - Automates key stages of systematic reviews including screening, data extraction and evidence synthesis. - Assists with ranking and prioritising studies.
ASReview	ML NLP	ASReview is an open-source, AI-assisted systematic review tool. - Uses active learning to accelerate screening, study retrieval and data analysis for systematic reviews. - Learns 'include' and 'exclude' decisions.
Elicit	LLM	Elicit is an AI-powered research assistant that supports systematic reviews with all outputs grounded in cited research sources. Capabilities include semantic search, screening, summaries data extraction, developing research questions and reports.

Reference management

Tool name	Type	Description
Zotero (ZotAI)	LLM	ZotAI is an AI research assistant that integrates with Zotero to help users work with PDF libraries. Uses context-aware AI models to organise, analyse and summarise PDFs in Zotero.

Tool name	Type	Description
EndNote 2025	Not AI, includes AI features (LLM)	EndNote is a reference management tool. Endnote 2025 includes AI features that help users evaluate and work with sources more efficiently. Generates summaries of PDFs and content based on PDF input, translates text, automates retraction alerts, automates journal matching.

Abbreviations: LLM, Large Language Model; ML, Machine Learning; NLP, Natural Language Processing.