Good Enough to Trust: Standards for Living Evidence and the Ethical Use of AI

Chair:

Firmaye Bogale, Public health Researcher at Ethiopian Public Health Institute

Speakers:

Ella Flemyng, Head of Editorial Policy & Research Integrity, Cochrane

Melanie Golob, Researcher in Living Evidence Synthesis & DPhil Candidate, University of Oxford

Brief Summary

This session of the Alive Learning Forum focused on standards for living evidence and the ethical use of AI in evidence synthesis. Hosted by Firmaye Bogale of the Ethiopian Public Health Institute, it featured two presentations: Ella Flemyng introduced RAISE, a shared framework for responsible AI use in evidence synthesis, and Melanie Golob shared findings from her research on whether existing standards for living systematic reviews are sufficient. A panel discussion followed on how to balance efficiency with rigour, how to decide when a review should be living, and how AI and publishing models shape what living evidence can achieve. The session closed with details of the next forum and a call for contributions to the Living Evidence Atlas.

Presentation by Ella Flemyng

Ella, Head of Editorial Policy and Research Integrity at Cochrane, introduced RAISE (Responsible AI Use in Evidence Synthesis), developed with colleagues from 30 organisations so that expectations for AI use do not fracture as the technology evolves. First released in September 2024, it is now a three-paper collection hosted on the Open Science Framework, with journal publication and a website expected soon.

The first paper sets out interlinked recommendations for eight roles in the evidence synthesis ecosystem, stressing that responsible AI use depends on everyone playing their part. The second covers building and evaluating tools, grounded in three principles: transparency, so AI contributions are explainable and reproducible; comprehensiveness, so AI does not introduce bias by favouring easily accessible research; and accountability, since responsibility never transfers to the tool. Tools should be built, tested on unseen data, and then validated for their intended task, with efficiency never coming at the expense of rigour.

The third paper addresses selecting and using tools, built around four expectations agreed by Cochrane, Campbell, JBI, and the Collaboration for Environmental Evidence: AI should be used

with human oversight, synthesists remain responsible for their research, tools must not compromise methodological rigour, and AI use must be fully reported. Its Responsible Handover Framework asks users to assess five domains before adopting a tool: its purpose, the source of its training and validation data, its validation and performance, the team's expertise and resources, and the transparency of its licences and documentation. Ella closed by emphasising that responsible AI is a shared responsibility, that synthesis principles come first, and that trust requires evidence rather than assumptions.

Presentation by Melanie Golob

Melanie, who manages a state-level health technology assessment programme in the United States and is completing her doctorate at the University of Oxford, explained that systematic reviews struggle to keep pace with primary research, with nearly 80 now published each day. Cochrane piloted living systematic reviews in 2017 and published standards in 2019, covering why a review should be living (decision priority, emerging evidence, and uncertainty) and how to conduct one, including search frequency, update thresholds, and criteria for ceasing updates.

Her 2022 study compared 23 Cochrane and 23 non-Cochrane living reviews against these eight standards. Only 4 of 46 met all eight, all of them Cochrane reviews. The least-reported item across both groups was when to stop updating, with common reasons including funding ending, a topic losing priority, or sufficient certainty being reached. A later update showed some improvement, but not substantial. Her key finding was that the "living" label alone is not enough: when readers cannot tell whether evidence is current, trustworthiness and policy usability suffer. Her Delphi process with experts has produced best practice statements, now under peer review and due to be presented in Sicily the following week.

Panel Discussion

On balancing efficiency with oversight, Ella acknowledged that completing the Responsible Handover Framework and verifying outputs currently adds to reviewers' workload. She urged teams to publish their evaluations so the field builds a cumulative evidence base and reduces duplication, and asked tool developers to make the information users need easy to find. On low- and middle-income contexts, she noted that validation data is often English-only and open access, underrepresenting LMIC research, so users should assess data sources carefully and apply mitigations suited to their review's context.

Melanie stressed transparency in deciding whether a review should be living: checking the topic against the three criteria, re-evaluating along the way, and stating upfront when updates will stop. She noted that COVID-19 showed how hard certainty is to reach when the topic itself keeps changing. On AI's role, Ella suggested living review authors may benefit most, since

validating a tool in the first version can streamline later ones. Melanie agreed AI is increasingly essential but identified the publication model as a structural barrier, with peer review timelines leaving updates outdated before publication and some reviews scattered across journals under different titles.

Responding to audience questions, Ella confirmed that RAISE is supported by Cochrane, Campbell, JBI, and the Collaboration for Environmental Evidence, though still a preprint, and that where tools are built on large language models developers do not control, confidence must rest on large-scale external validation. Melanie confirmed her Delphi panel included software and AI specialists.

Closing

The next Alive Learning Forum takes place on December 1st, focusing on living evidence in Africa and co-hosted by the Africa Evidence Network. Participants were also invited to submit any living evidence efforts missing from the Living Evidence Atlas, along with status updates on those already listed.

- The recordings of the session can be accessed on youtube: Here
- Our youtube channel : Here
- Our website: https://aliveevidence.org/
- To join the Alive Learning Forum, please sign up here.
- Any questions? Reach out at coordinator@aliveevidence.org
- The next session: **TBA**