

Artificial Intelligence (AI) in Oncology Health Technology Assessment (HTA): Current Applications, Limitations, and Implications for Living-HTA

Arup Pramanik¹, Stacy Grieve², Rozee Liu², Anna Forsythe²

¹Independent consultant, Boston, MA, USA; ²Oncoscope-AI, Miami, FL, USA

OBJECTIVES

- To review current AI applications across the oncology HTA evidence pathway, emerging guidance, and assessed improved efficiency from existing AI implementations for Living-HTA

BACKGROUND

- Oncology HTA increasingly requires complex, interconnected evidence packages spanning systematic literature reviews (SLRs), clinical study reports (CSRs), real-world evidence (RWE), and economic modelling.
- Growing evidence volumes and frequent updates created substantial operational bottlenecks. AI tools have been proposed to improve efficiency; however, concerns remain regarding validity, transparency, regulatory acceptance, and applicability to Living-HTA processes

METHODS

- A scoping review was conducted of peer-reviewed publications January 2023–December 2025, and HTA and regulatory guidance/methodological frameworks on AI-assisted evidence generation. We searched biomedical literature databases using terms on AI, machine learning, large-language models (LLMs), oncology, HTA, SLR, RWE, and economic modelling.
- Institutional websites and reference lists were reviewed.
- Selection focused on applications or guidance relevant to oncology HTA.
- Additional structured extraction was conducted to:
 - Map AI tool capabilities across SLR workflow steps
 - Assess coverage, integration, and output formats
 - Identify validation evidence and alignment with HTA requirements

RESULTS

- Across reviewed guidance, HTA bodies consistently emphasize transparency, validation, bias assessment, and human-in-the-loop governance as prerequisites for acceptability (Table 1)
- Detailed review of HTA and regulatory guidance confirms consistent expectations¹⁻⁸
 - Explicit disclosure of AI use in submissions
 - Validation against alternative or manual methods
 - Reproducibility and explainability
 - Bias identification and mitigation
 - Ongoing monitoring and governance
- These requirements reinforce that AI-assisted outputs must meet the same evidentiary standards as traditional approaches
- Most implementations remain fragmented, static, and project-based. Few support continuously-updated evidence synthesis or materially shorten end-to-end Living-HTA timelines, as outputs typically require re-execution when scope, comparators, or jurisdictional requirements change⁹⁻¹⁹ (Table 2)
- AI-assisted SLR tools were mapped across key workflow steps, including protocol design, search strategy, screening and review, data extraction, study-level synthesis, bias assessment, interpretation, and reporting outputs

Findings indicate

- Most tools support isolated steps rather than the full workflow
- Limited integration between screening, extraction, and synthesis
- Lack of standardized trial-level structuring and HTA-aligned outputs
- Outputs are typically static, project-based, and not designed for continuous updating

- Published validation studies across AI-assisted SLR approaches demonstrate screening sensitivity typically >95% under human oversight with the efficiency gains primarily at task level
- A review of recent (2025 – 2026) published validation studies evaluating integrated AI-assisted SLR systems report⁽²⁰⁻²²⁾; sensitivity approaching ~99%, specificity ~98%, and very low false-negative rates (<1%)
- These findings suggest that higher performance may be associated with integrated, continuously maintained systems rather than isolated task-level tools
- The structural limitation identified across tools is their reliance on project-based execution models, where SLRs are initiated per project, core steps (search, screening, extraction, synthesis) are repeated, and outputs are static at completion
- In contrast, emerging approaches described in the literature as Living SLR systems maintain²⁰⁻²¹ persistent SLR infrastructure, continuous evidence ingestion, and iterative updates across workflow steps
- This enables SLR transition from batch evidence processing to continuous evidence maintenance (Figure 1)
- Within this context, Real-Time AI-augmented Living SLR (REAL-SLR) approaches have been described as systems that:
 - Maintain core SLR processes continuously rather than per project
 - Integrate AI-assisted screening and extraction with ongoing updates
 - Support incremental evidence incorporation
 - Operate under human-in-the-loop validation frameworks
- These approaches align with HTA expectations, including transparency, reproducibility, and continuous evidence readiness

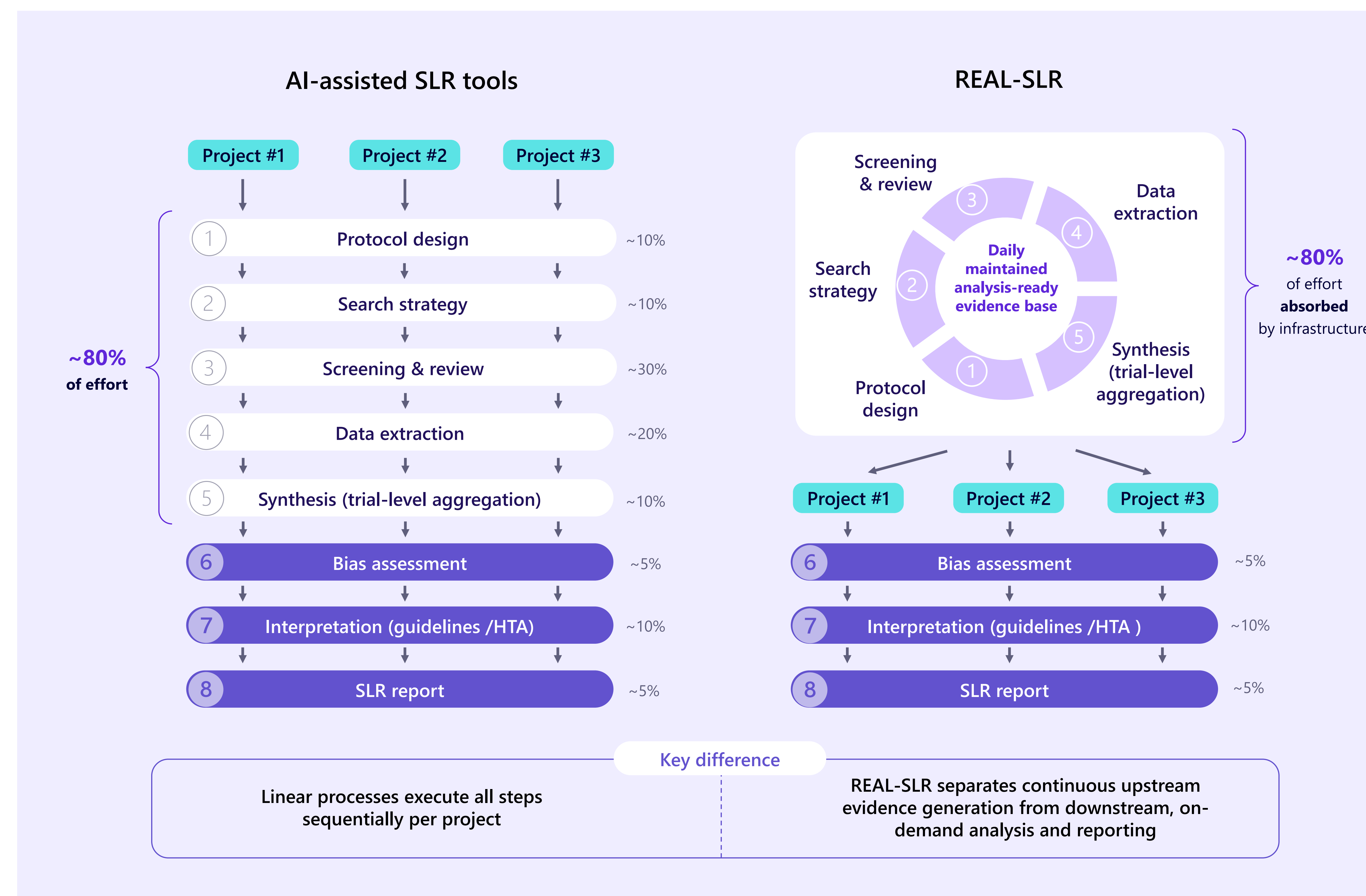
Table 1. Regulatory and methodological guidance on AI in evidence generation

Authority / Organization	Key Guidance Elements	Reference
NICE	<ul style="list-style-type: none"> Transparency of AI involvement required in submissions Validation against alternative methods; human oversight requirements Assessment of potential bias; reproducibility and explainability 	[1]
EMA	<ul style="list-style-type: none"> Documentation of AI use across the medicinal product lifecycle Risk management, validation, and compliance with data governance standards Labeling and disclosure of AI use to HTA bodies 	[2, 3]
FDA	<ul style="list-style-type: none"> Validation expectations for AI/ML systems; documentation of training data Risk management; transparency in regulatory submissions Ongoing monitoring post-approval 	[4]
Cochrane	<ul style="list-style-type: none"> Protocol registration of AI use in PROSPERO (systematic reviews) or OSF (scoping reviews) Validation studies demonstrating performance vs. manual approaches Bias assessment; reporting aligned with PRISMA extensions 	[5]
ISPOR	<ul style="list-style-type: none"> Methodological guidance on AI-assisted parameter extraction Validation requirements; comparison of AI-assisted vs. traditional methods CHEERS 2022 and modeling good practice guidelines for documentation 	[6, 7]
ISPE	<ul style="list-style-type: none"> Guidelines for AI use in pharmacoepidemiology Data governance and confidentiality requirements Bias assessment; validation and quality assurance standards 	[8]

Table 2. AI-assisted SLR tools and the SLR processes that they support

Tool	Protocol design	Search strategy	Screening & review	Data extraction	Synthesis (pub-trial grouping)	Bias assessment	Interpretation (guidelines / HTA context)	SLR report (Excel/Word)
DistillerSR ¹¹	! (protocol configuration possible but not a formal protocol authoring tool)	! (support, document, operationalize)	✓ (AI) Sensitivity: 14% ⁹ Specificity: 95% ⁹ False negative: 21% ⁹	✓ (AI)	! (can group via custom fields but no enforced trial hierarchy)	✓	X	! (CSV structured data download)
Covidence ¹²	X	X	✓ (AI) Uses most relevant sorting	✓	X	! (basic RoB templates; limited customization)	X	! (CSV structured data download)
EPPI-Reviewer ¹³	! (protocol configuration possible but not a formal protocol authoring tool)	X	✓ (AI) Uses priority screening	✓	X	✓	X	! (CSV structured data download)
Nested Knowledge ¹⁴	! (protocol configuration possible but not a formal protocol authoring tool)	! (support, document, operationalize)	✓ (AI) Sensitivity: 99% ¹⁴ Specificity: 46% ¹⁴ False negative: n/a	✓ (AI)	✓	✓ (AI)	! (supports visualization / dashboards but no direct HTA / guideline linkage)	✓
Rayyan ¹⁵	X	X	✓ (AI) Sensitivity: 78% ¹⁰ Specificity: 99% ¹⁰ False negative: 21% ¹⁰	X	X	X	X	! (reports available but limited formatting and reporting features)
Abstrackr ¹⁶	X	X	✓ (AI) Sensitivity: 60% ¹⁰ Specificity: 99% ¹⁰ False negative: 39% ¹⁰	X	X	X	X	X
ExaCT ¹⁷	X	X	X	✓ (AI)	X	X	X	! (exports extracted data but not full SLR reports)
RobotReviewer ¹⁸	X	X	X	X	X	✓ (AI)	X	! (outputs bias reports but not full SLR reporting formats)
LaserAI ¹⁹	X	X	✓ (AI) No published validation benchmarks	✓ (AI)	! (supports extraction relationship specification)	✓	X	! (AI) (generates summaries but not submission-ready reports)

Figure 1. AI-assisted SLR tools vs. living REAL-SLR framework



CONCLUSIONS

- While AI-assisted workflows show promise across components of oncology HTA, current implementations have rarely translated into sustained efficiency gains for Living-HTA.
- Achieving this objective will require approaches that move beyond task-level automation toward integrated, governed systems capable of continuous evidence maintenance.
- Clear implementation frameworks and governance models are needed before AI can reliably support living-HTA

REFERENCES

- National Institute for Health and Care Excellence. AI in Evidence Generation: Position Statement (ECD11). NICE; 2024.
- European Medicines Agency. Reflection Paper on the Use of Artificial Intelligence in the Medicinal Product Lifecycle. EMA; 2023.
- European Medicines Agency. Reflection paper on the use of artificial intelligence in the medicinal product lifecycle. Accessed May 4, 2026. <https://www.ema.europa.eu>
- US Food and Drug Administration. Real-World Evidence: Assessing Electronic Health Records and Claims Data to Support Regulatory Decision-Making. Guidance for Industry. FDA; 2023.
- Higgins JPT, Thomas J, eds. Cochrane Handbook for Systematic Reviews of Interventions. Version 6.4. Cochrane; 2023.
- Husereau D, Drummond M, Augustovski F, et al. CHEERS 2022 statement: updated consolidated health economic evaluation reporting standards. *Value Health*. 2022;25(1):10-31. doi:10.1016/j.jval.2021.10.008
- International Society for Pharmacoeconomics and Outcomes Research. ISPOR Health Economic Modeling Guidelines. ISPOR; 2023.
- Combi C, Amico B, Bellazzi R, et al. A manifesto on explainability for artificial intelligence in medicine. *Artif Intell Med*. 2022;131:102423.
- Gartlner G, Wagner G, Lux L, et al. Assessing the accuracy of machine-assisted abstract screening with DistillerAI: a user study. *Syst Rev*. 2019;8(1):277. doi:10.1186/s13643-019-1221-3
- Dos Reis AHS, de Oliveira ALM, Fritsch C, et al. Usefulness of machine learning software to screen titles of systematic reviews: a methodological study. *Syst Rev*. 2023;12(1):68. doi:10.1186/s13643-023-02231-3
- Distiller website. Accessed May 5, 2026 at <https://www.distillers.com/products/distiller-systematic-review-software>
- Covidence website. Accessed May 5, 2026 at <https://www.covidence.com/>
- EPPI-Reviewer website. Accessed May 5, 2026. <https://eppi.ec.ac.uk/cms/Default.aspx?tabid=3797>
- Nested knowledge documentation. Accessed May 5, 2026. <https://about.nested-knowledge.com/docs/autob/>
- Rayyan website. Accessed May 5, 2026. <https://www.rayyan.ai/>
- Abstrackr website. Accessed May 5, 2026. <https://abstrackr.com/>
- King's College London. Searching for Systematic Reviews & Evidence Synthesis: AI tools in evidence synthesis. Accessed May 5, 2026 at <https://libguides.kcl.ac.uk/systematicreview/>
- Marshall J, J, Kulper J, & Wallace B, C. (2016). RobotReviewer: evaluation of a system for automatically assessing bias in clinical trials. *Journal of the American Medical Association*. 316(16), 159-201. <https://doi.org/10.1093/jama.316.16.159>
- LaserAI documentation. Accessed May 5, 2026. <https://help.laser.ai/support/solutions>
- Sarkisian S, Liu R, Liu E, Forsythe A. A living systematic literature review for non-small-cell lung, prostate, and breast cancer built with an agentic text annotation system powered by large language models to assist treatment decision making. *J Clin Oncol*. 2025;43(16), suppl:13657. doi:10.1200/JCO.2025.43.16.suppl:13657
- Borghaei H, Forsythe A, Grieve S, et al. Living systematic review linked to guidelines and regulatory approvals as a treatment decision support tool in NSCLC. *J Thorac Oncol*. 2025;20(10 Suppl 1):S250-S251. Abstract P1.17.78.
- Liu R, Campden R, Liu EX, Correa G, Forsythe A. Validation of an agentic large language model system in the review stage of a real-time AI-assisted living systematic literature review: a solution to instant and easy access to clinical trial data. *Value Health*. 2026;29(Suppl 1):MSR18.

