

Ritesh Dubey<sup>1</sup>, Nicola Waddell<sup>2</sup>, Rajdeep Kaur<sup>1</sup>, Gagandeep Kaur<sup>1</sup>, Barinder Singh<sup>1</sup>, Mrinal Mayank<sup>1</sup>, Shubhram Pandey<sup>1</sup>  
<sup>1</sup>PharmacoEvidence, Mohali, India, <sup>2</sup>PharmacoEvidence, London, UK

## INTRODUCTION

- Systematic literature reviews (SLRs) are the gold standard for synthesizing clinical and health economic evidence, forming the foundation of health economics and outcomes research (HEOR) decision-making, health technology assessments (HTA), and regulatory submissions<sup>1</sup>
- Approximately 23% of all SLR studies become outdated within two years of publication, as reviewers fail to incorporate newly published evidence, highlighting the critical need for efficient, timely update mechanisms<sup>2</sup>
- Performing SLR updates demands significant effort due to the rapid growth of available evidence and the lack of detailed protocol documentation, making the update process particularly challenging when different researchers carry out the update<sup>3</sup>
- Retrieval-Augmented Generation (RAG) enhances large language models by integrating external knowledge retrieval to improve factual consistency and reduce hallucinations, offering particular promise for evidence synthesis applications in healthcare<sup>3</sup>
- Building on a previously conducted SLR (Kaur et al., 2025), this study evaluated a RAG-based multi-agent generative artificial intelligence (AI) framework to update existing SLR reports with newly identified evidence - preserving traceability, maintaining previously validated content, and ensuring continuous human oversight throughout the update process<sup>5,6</sup>

## OBJECTIVE

- To evaluate the accuracy, completeness, and evidence traceability of a RAG-based multi-agent generative AI framework in updating an existing SLR and its associated report with newly identified evidence, under continuous human expert oversight
- To assess the efficiency of the AI-assisted SLR update approach compared to conventional manual updating workflows, in terms of reduction in update timelines and human effort required to achieve publication-ready, regulatory-aligned outputs

## METHODS

- A previously conducted SLR and its associated report, originally developed in 2024, served as the base for this study, providing the foundational evidence structure, validated content, and established reporting framework for the update process<sup>5</sup>
- An AI-assisted SLR update was conducted by searching relevant databases to identify newly published studies over the two-year update period, supported by automated data extraction tools operating within a human-in-the-loop framework requiring approximately 10-20% human effort (Figure 1)
- Newly identified studies were incorporated into a RAG-based multi-agent framework, which updated relevant sections of the existing SLR report including humanistic burden, economic burden, and unmet needs without modifying previously validated content
- A standardized evaluation framework was implemented whereby all updated report sections were independently assessed by a human subject matter expert (SME) using a 5-point Likert scale across three domains: factual accuracy, completeness, and evidence traceability
- Iterative refinement was applied in cases where minor terminology adjustments or presentational improvements were required, ensuring all outputs met established SLR reporting standards before finalization

## RESULTS

- Over the two-year update period, five newly published studies were identified through the AI-assisted SLR update process, confirming the framework's ability to support a living SLR model, enabling continuous, periodic incorporation of newly published evidence into an existing validated report structure without requiring a full review and rebuild from scratch
- Compared with traditional manual updating workflows, the AI-enabled framework reduced update timelines by approximately 80-85%, significantly accelerating the evidence generation process and establishing a sustainable, resource-efficient update model for long-term living SLR maintenance in HEOR decision-making
- Automated data extraction tables were generated for all newly identified studies using a human-in-the-loop approach, requiring only approximately 10-20% human effort, substantially reducing the manual burden associated with conventional data extraction workflows while maintaining analytical accuracy
- The finalized extraction tables were subsequently uploaded into the framework and used to update the existing SLR report, the RAG-based architecture proved particularly suited to this living SLR workflow, retrieving and contextualizing only newly identified evidence against the existing report, ensuring updates were additive and targeted rather than requiring re-processing of the entire evidence base
- The framework generated revised narrative text, tables, and figures directly within the original manuscript structure without modifying previously validated content - meeting a critical requirement for living SLR compliance and ensuring continuity, traceability, and integrity of the existing report across update cycles
- All updated tables and figures were appropriately aligned with the evidence, and most updated report sections, including humanistic burden, economic burden, and unmet needs, were rated as "Strongly Agree" or "Agree" by the human expert on the 5-point Likert scale, reflecting high factual accuracy and clear linkage to source evidence. Minor terminology refinements were required in a limited number of cases, addressed through iterative human review
- The framework's ability to maintain evidence's traceability and audit-readiness across update cycles positions it as a viable tool for regulatory-aligned living evidence synthesis in HTA and reimbursement submissions

## CONCLUSIONS

- The RAG-based generative AI framework successfully supported a living SLR - enabling efficient, periodic incorporation of newly published evidence into an existing validated report structure, with update timelines reduced by approximately 80-85% compared to conventional manual workflows and requiring only 10-20% human effort for data extraction
- Continuous human expert oversight remained integral to ensuring terminology accuracy, methodological rigor, and adherence to SLR reporting standards - reinforcing that generative AI serves as a powerful enabler of responsible, efficient evidence synthesis rather than a replacement for specialist expertise

## References

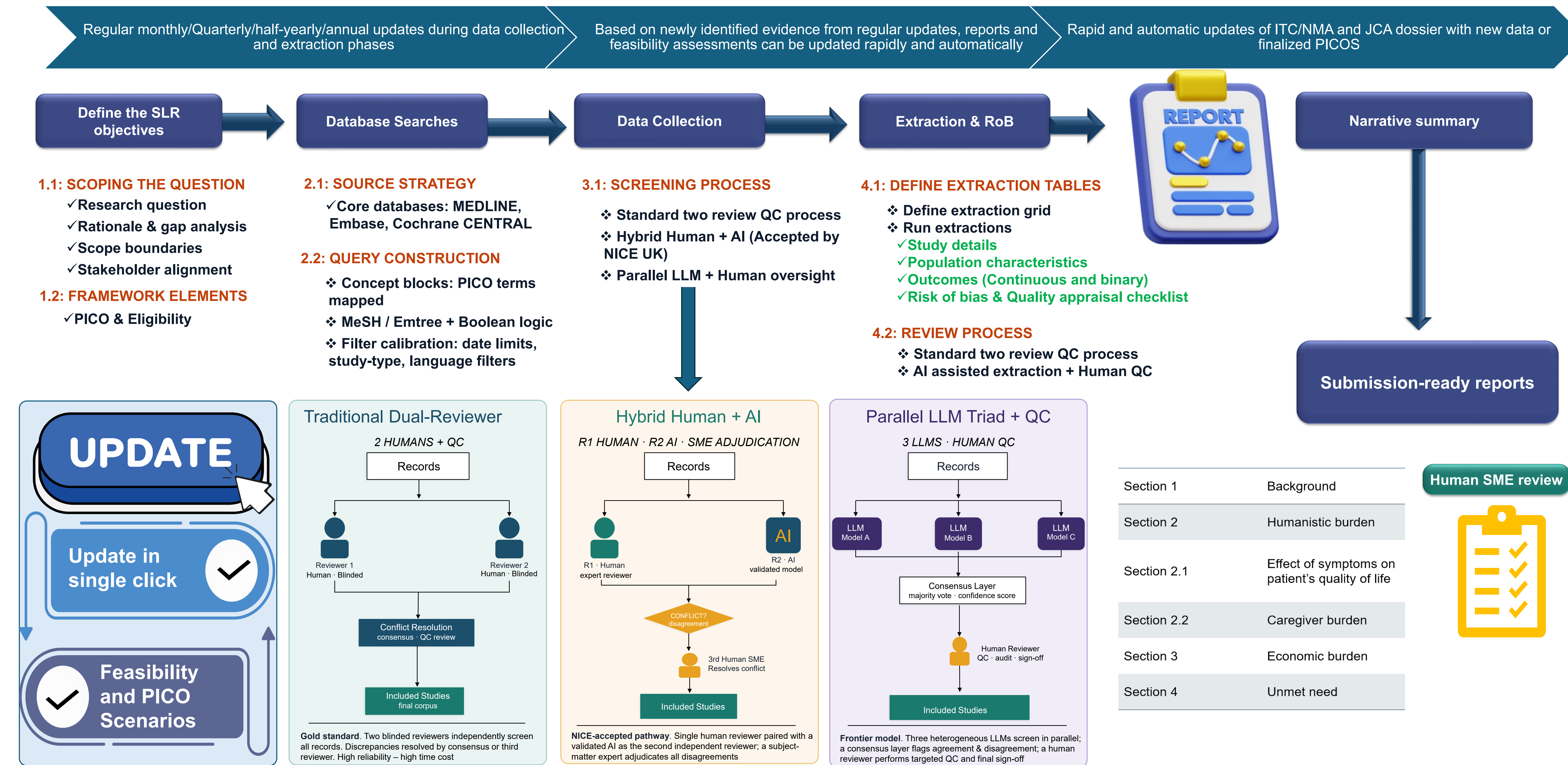
1. Grant MJ & Booth A. Health Inf Libr J. 2009;26(2):91-106; 2. van Dinter R et al. Inf Softw Technol. 2021. doi:10.1016/j.infsof.2021.108689; 3. Felizardo KR et al. arXiv. 2025. arXiv:2502.08050; 4. Zalka C et al. MDPI AI. 2025;6(9):226. doi:10.3390/ai6090226; 5. Kaur R et al. Value in Health. Volume 27, Issue S2, 2025; 6. Manion FJ et al. JMIR Med Inform. 2024. doi:10.2196/5465

Correspondence: Barinder Singh; barinder.singh@pharmacoEvidence.com

Acknowledgements: The authors wish to thank AK (Adarsh Kumar) & RA (Rupal Arora), for their valuable support in drafting this poster

Disclosure: SP, NW, RK, GK, BS, MM, and RD the authors declare that they have no conflict of interest

Figure 1: Schematic diagram of the AI-assisted SLR writing process



AI: Artificial Intelligence; LLM: Large Language Model; NICE: National Institute for Health and Care Excellence; PICO: Population, Intervention, Comparator, Outcome; QC: Quality Control; R1: Reviewer 1; R2: Reviewer 2; SME: Subject Matter Expert; UK: United Kingdom

## RESULTS (Continued)

- Efficient AI updating:** A RAG-based generative AI framework can efficiently update existing SLRs while maintaining transparency and full evidence traceability
- Human oversight:** Human-in-the-loop design ensures continuous oversight and compliance with evidence generation standards throughout the process
- Significant time savings:** Update timelines reduced by 80-85% compared to traditional manual workflows without compromising accuracy
- Significant cost efficiency:** Lower overall resource burden due to reduced human effort and streamlined workflows, compared to traditional manual approaches
- Scalable & compliant:** Offers a practical and responsible pathway for AI adoption in HEOR evidence generation at scale

## KEY FINDINGS

- Report sections rated "Strongly Agree" or "Agree" for factual accuracy
- Revised narrative text, tables, and figures generated within original manuscript structure
- Previously validated content preserved without modification
- Minor terminology refinements required in limited cases