View all Parexel's posters at ISPOR 2025

Artificial Intelligence-Powered Literature Reviews (AILRs): Streamlining Joint Clinical Assessment (JCA) Dossier Preparation for the Next Generation of Pharmaceutical Submissions

#MSR37

S. Budhia[1], G.S. Mangat[2], A. Jain[2], and S. Sharma[3]
Parexel International, [1]London, United Kingdom; [2]Mohali, India; [3]Chandigarh, India

Background

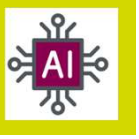
- The JCA, implemented in January 2025, creates a single harmonized clinical evidence evaluation for all 27 European Union (EU) member states. While JCA streamlines clinical assessments and promotes equitable access to innovative treatments, it presents significant challenges for pharmaceutical sponsors (Figure 1).
- Varying standards of care across EU member states may result in numerous Population, Intervention, Comparator, and Outcome (PICO) requests in the final scope, even after JCA consolidation. This requires sponsors to proactively anticipate JCA requirements.
- Additionally, JCA necessitates simultaneous evidence synthesis for all 27 member states without the traditional prioritization of key markets.
- With only 100 days between receiving the final scope and submitting the JCA dossier, sponsors face an extremely limited time for evidence generation, synthesis, and strategy development. This becomes even more complicated when sponsors must manage multiple stakeholders across member states, given the limited opportunities for direct engagement.

Figure 1: Areas of challenge for sponsors

- Comparative effectiveness evaluation forms the core of the JCA. While PICOS sets the framework, clinical systematic literature reviews (SLRs) provide the essential evidence foundation. However, traditional SLRs have their own set of limitations, particularly their complete reliance on human-only efforts, which becomes increasingly impractical within JCA's compressed timelines.
  - Artificial Intelligence (AI) offers a promising way forward, enabling more efficient evidence identification and extraction while maintaining the scientific rigor necessary for these submissions (Figure 2).

Figure 2: Potential advantages of using AI in clinical SLRs for JCA submission

- ✓ Accelerated Screening Process
- ✓ Enhanced Efficiency in Data Extraction
- ✓ Proactive Evidence Preparation
- ✓ Scalability for Multiple PICOS
- ✓ Consistency in Approach
- ✓ Resource Optimization
- ✓ Rapid Updates and Adaptability
- ✓ Improved Documentation
- ✓ Quality Control Automation



- Despite the numerous potential advantages of AI (Figure 2), its application has not been thoroughly explored in this context. Our study specifically focused on the screening and extraction steps of a clinical SLR, where we directly compared the performance of an AI-powered literature review (AILR) against traditional human-only approaches (Figure 2).

Objectives

- To evaluate the efficiency of AILR compared to traditional (human-only) methods in addressing JCA time constraints.
- To assess the accuracy of AI across three key phases: citation screening, full-text retrieval/screening, and data extraction/critical appraisal.
- To quantify time savings achieved through AILR implementation in systematic review processes relevant to JCA submissions.

Methodology

- Our study employed a three-phase comparative approach to evaluate the efficiency and accuracy of AILR against traditional human-only methods. As detailed in Table 1, we systematically assessed both approaches across the complete literature review workflow - from initial citation screening through data extraction and critical appraisal. Each phase incorporated appropriate quality control measures to ensure methodological rigor while enabling direct comparison between approaches.

Table 1: Comparative methodology – AILR vs. Human-only literature review approach

Phase	Process Steps	Human-Only Approach	AILR	Metrics Compared
1 Title/ Abstract Screening	• Systematic search: EMBASE & PubMed (n=4,124 citations) • Population: Adults with a neuroendocrine tumor • Focus: Controlled trials of systemic therapies	• 100% human screening by experienced reviewer (8+ years) • 30% quality check by second reviewer	• 30% human review for training of the AI model • 70% AI screening •Quality check for over-inclusive selection	• Accuracy • Precision • Recall • F1 score • Time efficiency
2 Full-Text Retrieval & Screening	• PICOS-based evaluation • Detailed assessment of retrieved articles (n=558 citations)	• 100% human screening by the same reviewer from phase 1 • 30% quality check by second reviewer	• General automated technology (non-AI) features: Bulk PDF downloading - Automatic tagging – AutoPRISMA	• Time efficiency
3 Data Extraction & Critical Appraisal	• Extraction of key outcomes • Quality assessment (n=20 studies)	• Human extraction • Critical appraisal using industry standard tool	• LLM with standardized prompts • Structured extraction of: Median OS - Median PFS - Response rates • Comprehensive critical appraisal	• Accuracy • Time efficiency

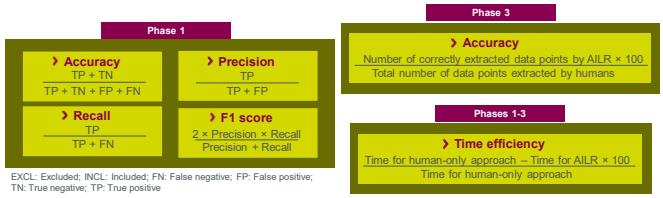
LLM: Large language model; OS: Overall survival; PFS: Progression-free survival; PICOS: Patient population, intervention, comparator, outcomes, study design

Methodology….

- Our study design deliberately accounted for varying maturity levels of AI applications across different review phases.
  - For title/abstract screening, where AI capabilities are most advanced, we implemented comprehensive machine learning approaches with extensive performance metrics.
  - For full-text screening, where AI applications remain less established, our methodology incorporated automated efficiency-enhancing features (bulk PDF downloading, automatic tagging) rather than direct AI automation.
  - In the data extraction phase, we strategically limited our scope to three standardized clinical outcomes (OS, PFS, response rates) to align with the current developmental stage of extraction AI algorithms.

This measured, pragmatic implementation across phases allowed for a realistic assessment of AI's current capabilities while establishing a foundation for future scalability as technologies mature.

Figure 3: Evaluation framework and performance metrics for AILR vs. Human-only approach

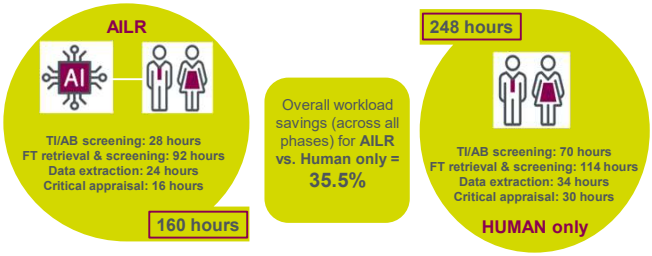


EXCL: Excluded; INCL: Included; FN: False negative; FP: False positive; TN: True negative; TP: True positive

Results

- The AILR approach demonstrated substantial time savings across all review phases, with the most pronounced efficiency gains observed in the initial screening phase (Figure 4).
- Title/abstract screening showed the highest relative time efficiency improvement, highlighting AI's maturity and effectiveness in this well-established application area.
- Despite the more limited AI implementation in full-text retrieval and screening, even the integration of automated features yielded meaningful time savings compared to fully manual processes.
- Data extraction and critical appraisal, despite being more complex cognitive tasks, showed impressive efficiency improvements when supported by AI, suggesting significant potential for future applications.
- The cumulative time savings across all phases demonstrates that even selective implementation of AI in a literature review workflow can substantially reduce resource requirements (Figure 4).

Figure 4: Comparative time-savings analysis (AILR vs. Human only)



AB: Abstracts; FT: Full-texts; TI: Titles

- Time efficiency improvements were achieved while maintaining robust performance metrics in:
  - screening of titles/abstracts (92.1% accuracy, 61.19% precision, 91.03% recall, 73.20% F1 score),
  - data extraction (91.3% accuracy), and
  - critical appraisal (88.0% accuracy), demonstrating AI's capability to address JCA's compressed timelines without significantly compromising evidence quality.

Conclusion

- Based on our findings, the AILR approach delivers significant advantages for navigating the JCA process. We can deliver early evidence assessment, mapping the sponsor's complete evidence landscape well before the critical 100-day JCA period begins. This early insight allows for identification of potential evidence gaps when there's still time to address them.
- Our hybrid methodology combines AI efficiency in title/abstract screening with human expertise for complex analysis. This approach ensures both speed and accuracy, producing documentation that withstands rigorous JCA and HTA scrutiny.
- The efficiency gains from our AILR approach translate directly into strategic advantages for the sponsor's submission. By accelerating the evidence synthesis process, we will create additional time for refining value narratives and strengthening dossier elements, ultimately enhancing the product's positioning in the European HTA landscape.