DistillerSR

How Eli Lilly Reduced Screening Burden by 67% With DistillerSR AI

A real-world evaluation of DistillerAI and Classifiers across 11 targeted and systematic literature reviews

67% Reduction in Screening Burden
Across 11 literature reviews, AI reduced overall human effort by two thirds – 77.9% in TLRs and 57.3% in SLRs.

90% Median Accuracy Classifiers and DistillerAI achieved 90% median accuracy (range: 85–96%) with only 1.64% mean false-negatives – comparable to manual screening.

Reusable Training Sets
Training sets built in one review can be reused across reviews with similar scope, compounding efficiency gains over time.

AI Partially Replaces Second Reviewer
In SLRs, Classifiers can screen the test set in place of a second reviewer, reducing resource requirements without sacrificing rigour.

> "AI is found to be an efficient tool for title-abstract screening, especially for large reference sets. AI simulation tools are useful in prioritising likely inclusions and exclusions; however, additional quality checks are required to meet the rigorous requirements of HTA."
>
> **Venkata et al., Eli Lilly and Company I ISPOR European Congress, Copenhagen, 2023**

The Challenge: Screening at Scale

Systematic and targeted literature reviews are foundational to drug development and HTA submissions – but they are resource-intensive at every stage. Title-abstract screening alone, when conducted manually across thousands of citations, places significant demands on reviewer time and introduces the risk of human fatigue and inconsistency.

For the clinical evidence teams at Eli Lilly, managing literature reviews across a broad therapeutic portfolio meant that screening efficiency was not just a workflow preference – it was a strategic requirement. The question was whether AI-assisted tools could meaningfully reduce that burden without compromising the accuracy and rigour that HTA bodies demand.

The Approach: DistillerSR AI and Classifiers

Between February 2021 and June 2023, Eli Lilly evaluated DistillerSR's two AI features across 11 literature reviews – eight targeted literature reviews (TLRs) using DistillerAI, and three systematic literature reviews (SLRs) using Classifiers.

For TLRs, DistillerAI applied a naïve Bayesian approach: a training set of 10–20% of manually screened citations was used to assign likelihood-of-relevance scores (0 to 1) to the remaining references, enabling reviewers to prioritise likely inclusions and exclusions efficiently.

For SLRs, Classifiers – DistillerSR's advanced AI engine – used natural language processing to screen all unreviewed citations in one of the two-reviewer sets. Performance was validated using balance and recall scores, giving reviewers confidence in the outputs before accepting them.

In both cases, human review of AI outputs was retained as a mandatory quality check – a deliberate design decision that maintained audit trail integrity and regulatory credibility throughout.

The Results: Substantial, Measurable Time Savings

Across all 11 reviews, the use of AI reduced overall human screening effort by approximately 67% – a figure that held across both review types, with TLRs achieving 77.9% efficiency gains and SLRs achieving 57.3%.

For TLRs, the total citations screened across eight reviews numbered 17,470. Without AI, the estimated screening time was 399 hours; with AI, that figure dropped to 88 hours of AI training time, saving 311 hours of reviewer effort. For the three SLRs covering 20,100 citations, AI saved 263 hours against a baseline of 459 hours.

Accuracy remained high throughout. The median accuracy score across all reviews was 90%, with a range of 85–96%. Mean false-negatives – the most consequential error type in systematic reviews – stood at just 1.64%, a rate comparable to manual screening performance.

The lower accuracy observed in SLRs (85%) compared to TLRs (94%) was attributable to two factors: the smaller sample size of SLR projects tested, and a possible unequal distribution of inclusions and exclusions in the training set – both identified as areas for further refinement.

Looking Ahead

The Eli Lilly evaluation demonstrates that AI-assisted screening in DistillerSR can substantially reduce the front-end burden of literature reviews without compromising accuracy – provided that quality checks remain part of the workflow.

As HTA bodies including NICE and CDA begin issuing guidance on the use of AI in evidence generation, the model demonstrated here – AI acceleration with documented human verification – aligns closely with emerging best practice for defensible, audit-ready submissions.

Further research to establish recommendations for optimal AI integration in SLRs for HTA submissions remains an active area of development, with DistillerSR continuing to evolve its AI capabilities in line with regulatory expectations.

About Eli Lilly and Company

Eli Lilly and Company is a global pharmaceutical leader headquartered in Indianapolis, Indiana, with operations in more than 125 countries. The clinical evidence teams at Lilly are responsible for conducting rigorous literature reviews in support of drug development and regulatory submissions worldwide.

About DistillerSR

DistillerSR is the market leader in AI-enabled literature review automation and evidence management, trusted by more than 80% of the world's largest pharmaceutical and medical device companies. Its platform supports defensible, audit-ready evidence management at scale.

Source: Venkata SK et al. Application of Artificial Intelligence in Literature Reviews. Poster MSR99. ISPOR 26th Annual European Congress; Copenhagen, Denmark; 12–15 November 2023. Sponsored by Eli Lilly and Company.