Mapping Human Effort Displacement in AI-Assisted HEOR Workflows

MSR44

Hanan Irfan, MSc[1], **Tushar Srivastava, MSc[1]**, Shilpi Swami, MSc[1]
[1]ConnectHEOR, London, UK.
Email: tushar.srivastava@connectheor.com

connectHEOR

KEY MESSAGE

AI doesn't remove human effort in HEOR - it redistributes it. Less time executing, more time judging, validating, and governing.
AI reduces procedural execution by 40–60% in HEOR workflows, but new demands — validation, governance, prompt management shift the bottleneck from analyst execution to senior expert judgement.

BACKGROUND

The integration of artificial intelligence (AI) into Health Economics and Outcomes Research (HEOR) workflows alters how human effort is mostly allocated across evidence generation and decision-support activities.[1]

Rather than functioning solely as a net reducer of workload, AI-assisted tools may displace effort toward oversight, governance, and accountability tasks.

This shift is increasingly reflected in regulatory expectations: recent guidance from the National Institute of Clinical Excellence (NICE) and Canada's Drug Agency states that AI use in evidence generation must be judged not only on efficiency, but also on appropriateness, explainability, and the ability to maintain accountability for submitted evidence.[2].

This study aimed to quantify and characterize how human time, cognitive load, and expert responsibility are redistributed across HEOR workflows following AI adoption.

OBJECTIVES

1. Quantify how human time is redistributed across HEOR workflows following AI adoption
2. Characterize shifts in cognitive load and required expertise levels
3. Map newly introduced task categories (prompt engineering, validation, audit trails) vis-a-vis human effort required
4. Compare displacement patterns across exploratory, internal decision-support, and HTA-facing use cases

METHODS

Study Design

A comparative workflow analysis contrasting traditional and AI-assisted processes across five HEOR domains was undertaken:

- SLR Support
- Evidence Synthesis
- Economic Modelling
- Dossier Drafting
- Quality Assurance (QA)

Four Assessment Dimensions

- Time Allocation
- Cognitive Load
- Expertise Level
- Account-ability

Human effort was assessed across four dimensions: time allocation, cognitive load, required expertise level, and accountability. Governance burden were stratified by use case, including exploratory analyses, internal decision support, and HTA-facing deliverables.

RESULTS

40–60% Reduction in Procedural Execution Time

~15% Workflow Time spent on Prompt Engineering

Cognitive Load Increase in QA & Validation

Shift From Content Production to Adjudication

Time Spent Across Task Categories

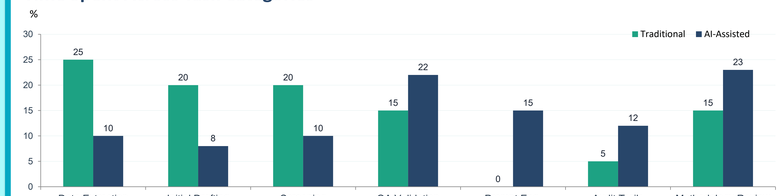


Fig 1: Time spent (in hours) across task categories via traditional vs AI assisted workflow

Governance Burden by Use Case

Exploratory Analyses — LOW GOVERNANCE
- Lower governance burden
- Higher tolerance for iterative refinement
- Formal traceability requirements lower

→ **Internal Decision Support** — MODERATE GOVERNANCE
- Moderate governance
- Reduced turnaround time
- Relatively lower documentation

→ **HTA-Facing Deliverables** — HIGH GOVERNANCE
- Highest governance burden
- Senior review & methodological arbitration
- Full traceability & accountability required

QA Focus Shift: Traditional vs AI-Assisted

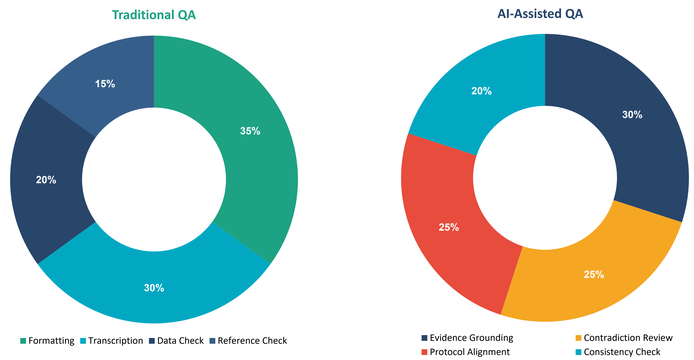


Fig 1 Time spent in QA of outputs across various categories in traditional vs AI assisted workflow

As AI takes over more procedural output generation, quality assessment (QA) effort is displaced from simple error detection to adjudicating outputs.

DISCUSSION

Key Findings

The redistribution pattern is consistent with published evidence showing that AI systems can accelerate screening, extraction, and evidence synthesis subtasks, but still require human-in-the-loop monitoring, prompt adjustment, and verification of intermediate outputs.

Recent empirical studies in AI-assisted review workflows have reported meaningful reductions in screening and extraction time together with retained or improved performance when experts remain actively involved.

Time released from repetitive execution appears to be reabsorbed by supervision-intensive activities, particularly where outputs must be submission-ready, reproducible, and defensible.

In QA, the task focus moved from detecting simple formatting or transcription issues to adjudicating whether AI-generated content was evidentially grounded, internally consistent, and aligned with protocol or modelling intent [3].

Implementation Implications

Implementation planning for AI in HEOR should be based on role redesign rather than assumptions of straightforward labor substitution or time savings.

The main operational bottleneck may shift from analyst execution time to the availability of experienced reviewers who can judge fitness for purpose, resolve ambiguity, and certify traceability.

AI systems can accelerate screening, extraction, and evidence synthesis subtasks, but still require human-in-the-loop monitoring, prompt adjustment, and verification of intermediate outputs.

CONCLUSIONS

1. AI-assisted HEOR workflows primarily reallocate, rather than eliminate, human effort. Efficiency gains are most reliable when AI supports the generation of candidate outputs, while humans retain responsibility for interpretation, methodological judgement, validation, and accountability.
2. The operational bottleneck shifts from analyst execution time to availability of experienced reviewers capable of judging fitness for purpose, resolving ambiguity, and certifying traceability across the workflow.
3. The value of AI assistance in HEOR depends not only on model capability, but also on governance design, documentation discipline, and the maturity of human-in-the-loop review structures required to produce credible evidence for internal and external decision makers.

References

1. Srivastava T, Irfan H, Babiy V, Swami S. Integration of generative AI with human expertise in HEOR: a hybrid intelligence framework. *Adv Ther*. 2025 Sep;42(9):4103-4130. doi:10.1007/s12325-025-03273-w.
2. National Institute for Health and Care Excellence (NICE). *Use of AI in evidence generation: NICE position statement* (ECD11). Published 15 August 2024.
3. Harmsen W, et al. *Machine learning to optimize literature screening in medical guideline development. Systematic Reviews.* 2024;13:177. doi:10.1186/s13643-024-02590-5.

Financial Disclosure:
Authors are employees of ConnectHEOR Limited. No external funding received to conduct this research and no conflict of interest to declare.

Poster presented at ISPOR in Pennsylvania, Philadelphia, USA, May 17-20, 2026