MSR207

Evaluating an AI-assisted Workflow for Table Data Extraction in HEOR Systematic Reviews

Ewa Borowiack, Monika Opalek, Ewelina Sadowska, Artur J. Nowak
Evidence Prime, Krakow, Poland

Keypoints

- First study to quantify AI-supported data extraction from HEOR tables.
- Laser AI reduced extraction time by ~50% while maintaining comparable accuracy to manual methods

Reduction of median extraction time per table	50%
Absolute time savings per table	10:15 min
Per-cell extraction time improvement	50%

- Senior reviewers with HEOR expertise achieved the largest time savings with Laser AI (71.9% faster than Excel).
- Error patterns were table-dependent rather than tool-dependent.

Objectives

This study aimed to quantify the time savings achieved by using the Laser AI workflow for extracting tables, compared with the conventional manual extraction process in Excel during Health Economics and Outcomes Research (HEOR) systematic reviews. The study also aimed to assess the frequency and types of errors occuring during extraction with both methods .

Introduction

HEOR generates a rapidly growing body of evidence to inform reimbursement and policy decisions. These reports are increasingly produced under tight timelines and there is a growing demand for creating rapid reviews. Much of the key evidence is reported in complex tables, extraction of which is error-prone and time consuming process because of the density of numerical values and the risk of transcription errors, despite quality-control checks.

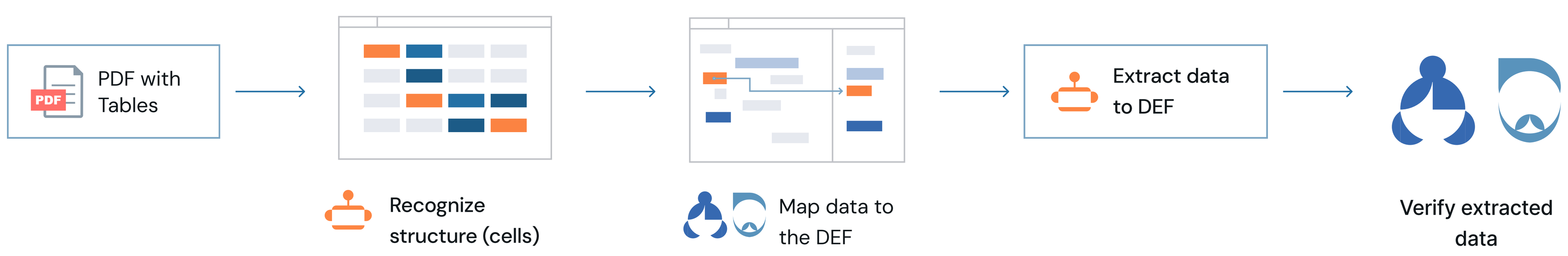
Artificial intelligence (AI) offers an opportunity to accelerate and streamline review process by automating repetitive manual tasks in systematic reviews. AI tools are increasingly accepted and adopted in systematic reviews, mainly during the searching and study selection [1]. Moreover, national HTA agencies have already started to use AI in literature reviews and have formulated guidelines. [2,3]

One aspect of the review that still raises questions regarding the use of AI is data extraction. Currently, AI-supported data extraction is not considered to be reliable support [1]. Therefore there is a growing number of solutions being tested to improve the quality of data extraction, including data from tables, which contain a lot of HEOR data.

Laser AI table extraction module uses vision models to analyze the document, find the table location and extract the table's structure and content. The workflow for semi-automated approach for table extraction consists of several stages [Fig 1]:

1. AI-driven table recognition
2. Human mapping to the Data Extraction Form (DEF)
3. AI-driven extraction into the DEF
4. Human validation

[Fig 1]

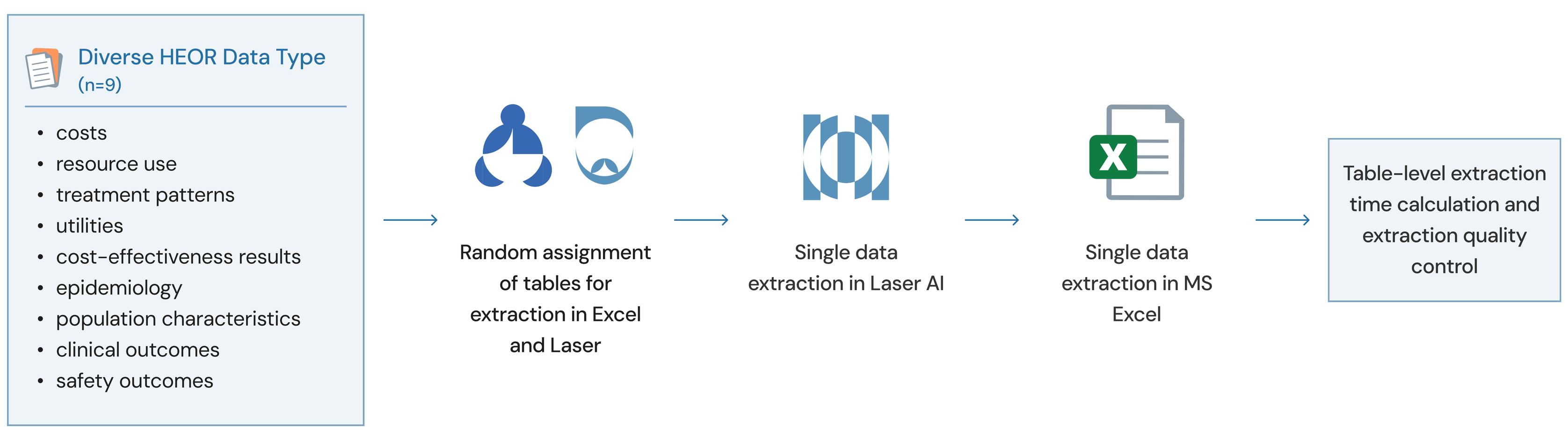


Methods

Nine tables representing diverse HEOR-related data types: (costs, resource use, treatment patterns, utilities, cost-effectiveness results, epidemiology, population characteristics and transition probabilities for clinical and safety outcomes [4-12]) were purposefully sampled from published studies. Three reviewers of differing experience extracted each table once with Laser AI and, in a separate session, once with Microsoft Excel. Table-reviewer assignments were alternated to avoid learning effects. Table-level extraction time was recorded. Per-cell extraction speed (seconds/value) was calculated post-hoc by dividing table time by value count. To determine whether AI-supported extraction performance varied according to domain knowledge and tool experience we conducted a subgroup analysis based on two independent factors [Fig 2]: HEOR domain expertise (Senior vs Junior), Tool familiarity (Expert vs Junior).

After the extraction stage, quality assurance (QA) of the extracted data was also performed. For Laser AI, a dedicated QA module was used, while in Excel the results extracted to the sheet were compared with the data in the original publication.

[Fig 2]

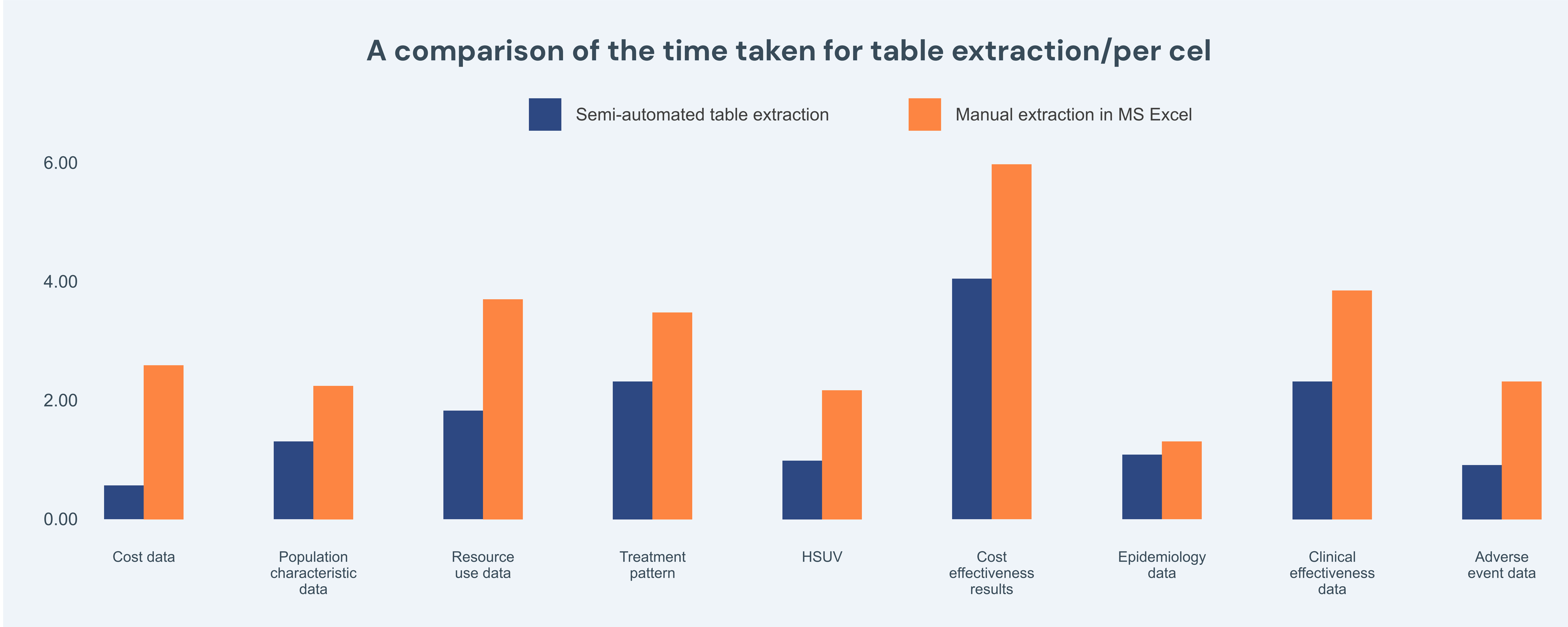


Results

Main Analysis

The nine tables contained 5 058 discrete values (median 378; range 112 - 1995). Median extraction time decreased from 21:43 min/table (7:10 – 77:25) with Excel to 10:58 min/table (4:19 – 43:38) using Laser AI, yielding a 50% relative reduction and an absolute saving of 10:15 min/table. Estimated per-cell speed improved from 2.6 s to 1.3 s (50%). Figure 3 presents detailed time data corresponding to each table type. Because timing was captured only at table level, per-cell metrics are illustrative rather than directly measured. Observed savings were consistent across reviewers despite divergent familiarity with the software. Completeness of cell extraction was 100% in both workflows.

[Fig 3]



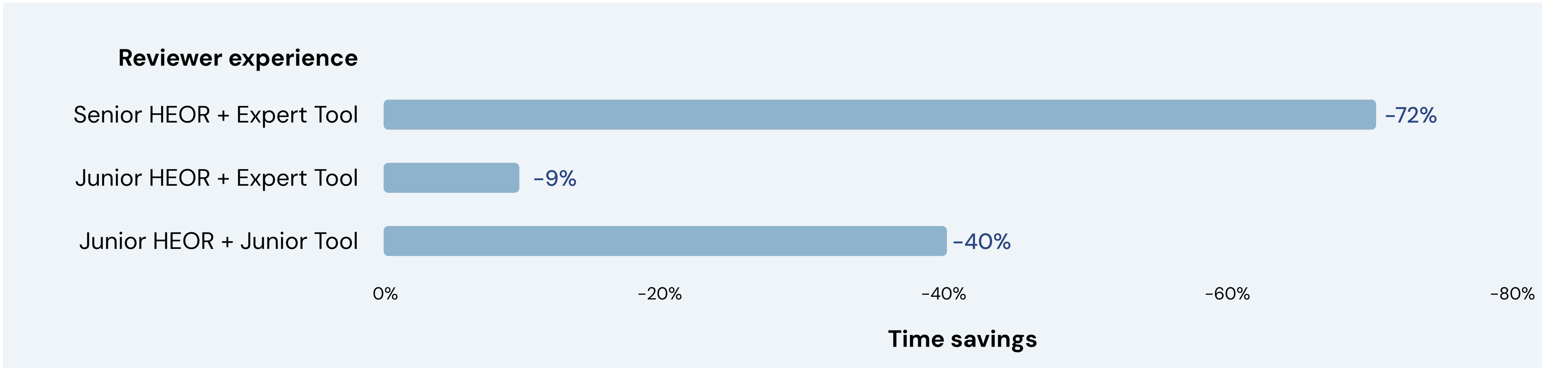
Reviewer Subgroup Analysis

We stratified results by both HEOR expertise and familiarity with the extraction tool [Fig 4]:

- Senior HEOR + Expert Tool users achieved the largest efficiency gain (−72%).
- Junior HEOR + Expert Tool users showed minimal improvement (−9%).
- Junior HEOR + Junior Tool users still benefited substantially (−40%).

This indicates that tool expertise is a major factor in addition to domain expertise. AI extraction offers the greatest time savings when reviewers are both domain experts and skilled in tool use, but even less-experienced reviewers see significant benefit.

Time saving vs. Rewiever experience [Fig 4]



QA review analysis

A QA review of the extracted values revealed errors in the data extracted using both Laser AI and Microsoft Excel. In the Becerra 2015 table[5], missing minus signs for negative values were observed in both Laser AI (five instances) and Excel (two instances). In the largest table (Boeckxstaens 2024[6]), typographical errors were detected in both workflows; additionally, in Laser AI extraction one data row was omitted and values from two adjacent rows were misaligned into incorrect fields. Similarly, during Excel extraction, one data row was omitted in the Nieto 2022[8] table, and in the King 2024[9] table values from one row were shifted upward by one row. Interestingly, error patterns appeared to be driven more by the complexity of the table itself than by the extraction tool.

Limitation

- Small sample size - the study only used 9 tables (one table per each data type) which may not be a large enough sample to generalize the findings to all types of tables in HEOR systematic review.
- Due to small sample size (9 tabels) only qualitative QA was performed - This experiment doesn't quantify or systematically analyze the specific features of table complexity that contribute to errors, making it difficult to pinpoint the exact causes.
- Given the small sample size and small number of reviewers, subgroup differences should be interpreted with caution; the relative impact of HEOR expertise and tool familiarity on efficiency gains cannot be reliably quantified.

Conclusion

This is the first known project to evaluate how AI can support data extraction from tables in HEOR reviews.

An AI-assisted extraction halved the time required to collect tabular HEOR data while completeness and error rates were comparable to the manual approach, demonstrating clear operational benefits for systematic review teams.

Future work will audit accuracy and explore learning-curves. As we transition toward a fully automated extraction process, we anticipate further improvements in time savings and efficiency.

References

1. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.5 (updated August 2024). Cochrane, 2024. Available from www.cochrane.org/handbook. [accessed 14.10.2025]
2. Use of AI in evidence generation: NICE position statement. https://www.nice.org.uk/about/what-we-do/our-researchwork/use-of-ai-in-evidence-generation--nice-position-statement [accessed 14.10.2025]
3. Canada's Drug Agency Position Statement on the Use of Artificial Intelligence in the Generation and Reporting of Evidence. https://www.cda-amc.ca/sites/default/files/MG%20Methods/Position_Statement_AI_Renumbered.pdf [accessed 14.10.2025]
4. Almatroudi, A. (2022). The Incidence Rate of Esophageal Cancer in Saudi Arabia: An Observational and a Descriptive Epidemiological Analyses.
5. Becerra, V., A. Gracia, K. Desai, S. Abogunrin, S. Brand, R. Chapman, F. G. Alonso, V. Fuster and G. Sanz (2015). Cost-effectiveness and public health benefit of secondary cardiovascular disease prevention from improved adherence using a polypill in the UK.
6. Boeckxstaens, G., S. Ayad, G. Dukes, M. Essandoh, R. Gryder, P. Kamble, J. Tackett, P. Thakker, J. Williams, Y. Zhang and P. R. Wade (2024). A randomized phase 2 study of the 5-HT 4 receptor agonist felcisetrag for postoperative gastrointestinal dysfunction after bowel surgery.
7. Healey, M. J., S. N. Brian, D. Princic, Black, M.-M. Elisabetta, S. Nilofer, Azad, L. Rory, Smoot, N. Princic, Á. D. Black, Á. E. Malangone-Monaco, N. S. Azad and R. L. Smoot (2022). Real-World Analysis of Treatment Patterns, Healthcare Utilization, Costs, and Mortality Among People with Biliary Tract Cancers in the USA.
8. Jesus Rodríguez-Nieto, M., E. Cano-Jiménez, A. D. Romero Ortiz, A. Villar, M. Morros, A. Ramon and S. Armengol (2023). Economic Burden of Idiopathic Pulmonary Fibrosis in Spain: A Prospective Real-World Data Study (OASIS Study).
9. King, B., M. M. Senna, N. A. Mesinkovska, C. Lynde, M. Zirwas, C. Maari, V. H. Prajapati, S. Sapra, P. Brzewski, L. Osman, S. Hanna, M. C. Wiseman, C. Hamilton and J. Cassella (2024). Efficacy and safety of deuruxolitinib, an oral selective Janus kinase inhibitor, in adults with alopecia areata: Results from the Phase 3 randomized, controlled trial (THRIVE-AA1).
10. Lebwohl, M., M. Bukhalo, L. S. Gold, B. Glick, M. Llamas-Velasco, S. Sanchez-Rivera, A. Pan, T. Zhan, L. Drogaris, K. Douglas, G. St, John, R. Espaillat and R. Bissonnette (2024). A randomized phase 3b study evaluating the safety and efficacy of risankizumab in adult patients with moderate-tosevere plaque psoriasis with nonpustular palmoplantar involvement.
11. Pickard, A. S., Y. Yang and T. A. Lee (2011). Comparison of health-related quality of life measures in chronic obstructive pulmonary disease.
12. Yoshida, M., N. Taguchi, Y. Piao, R. Gupta, M. Berry, J. Peters, M. Abdelghany, M. Chiang, C.-Y. Wang and H. Yotsuyanagi (2024). Treatment pattern and clinical outcomes of remdesivir in hospitalized COVID-19 patients with severe chronic kidney disease: a database analysis of acute care hospitals in Japan.