A targeted literature review on Artificial Intelligence and Large Language Models used to build Health Economics models

⋙ J.E. Poirrier [1], R. Bergemann [2]

[1] Parexel International, Wavre, Belgium

[2] Parexel International, Basel, Switzerland

Background

With the advent of Large Language Models (LLMs, like OpenAI's GPT-4 [1]), Healthcare, Access and Health Economics (HE) were not spared by the hype around Artificial Intelligence (AI). Several posters, panels and peer-reviewed publications tout a bright future for AI in building HE models [2, 3, 4]. Some try to present a balanced view that shows potential biases and issues with ethics, morality, privacy, fairness, responsibility, etc. [5, 6, 7].

Objective

This research aims to find concrete examples of HE models built using AI (or LLMs) and to analyse the authors' approach to pragmatic and ethical issues.

Methods

A targeted literature review (TLR) was performed to identify available cost-effectiveness models built by AI (or LLM). Electronic searches were conducted in PubMed (for published papers), EconPapers, ArXiv and Preprints.org (for preprints) to identify studies published from inception to the end of February 2024. The search was conducted using ("artificial intelligence" OR "LLM" OR "GPT-4") AND ("health economics" OR "cost-effectiveness") in titles and keywords to retrieve potentially relevant publications. The search procedure was completed manually. Publications were screened for inclusion and followed the population, interventions, comparators, and outcomes framework eligibility criteria in Table 1. In our review, the articles were included if they described a health economics evaluation. Studies were excluded (1) if AI and related techniques like LLMs were not used to generate the model, (2) if they were not published in English, (3) if their full text was not available, and (4) if they did not conduct an original economic evaluation. Potential duplicate articles were removed. One reviewer screened titles and abstracts. The same reviewer screened the full text. Quality assessment of the identified studies was planned for transparency, reusability and updatability of presented models [8].

Table 1. Population, Interventions, Comparators, and Outcomes

Criteria	Content
Population	All populations
Intervention	All interventions
Comparator	All comparators
Outcomes	All outcomes if the result of a health economics model (cost-effectiveness or budget impact) described in the publication

## Results

PubMed, EconPapers, ArXiv and Preprints.org returned 470, 5, 0, and 35 hits, respectively (total: 510 papers, Figure 1).

Among those, 270 papers (52%) described either a general perspective on AI in healthcare or a review of AI in healthcare (165 papers, 61%) or described an application of AI in healthcare (105 papers, 39%). 83 papers (16%) were about HE modelling, and among them, 50 papers (9%) described an HE model of an intervention including AI (either AI was the intervention or AI-augmented an existing intervention) (Figure 2).

Of all the HE modelling papers, only one (0.2%) described using AI to develop an HE model [9]. In this paper, authors prompted GPT-4 to program a copy in R of two existing cost-effectiveness models in oncology. Prompts and generated R code are published (transparency criteria from the Hamilton framework), but the copyrights are not defined (reusability criteria) and the code can't be reproduced (some functions and paths have been manually changed for the publication). The prompts feeding the LLM and the output of the LLM are shared, and no information about the LLM itself (state, version, corpus at the time of queries, …) and the query mechanism (API, code, version, endpoints, …) is given.

Figure 1. Flow chart of the study selection process
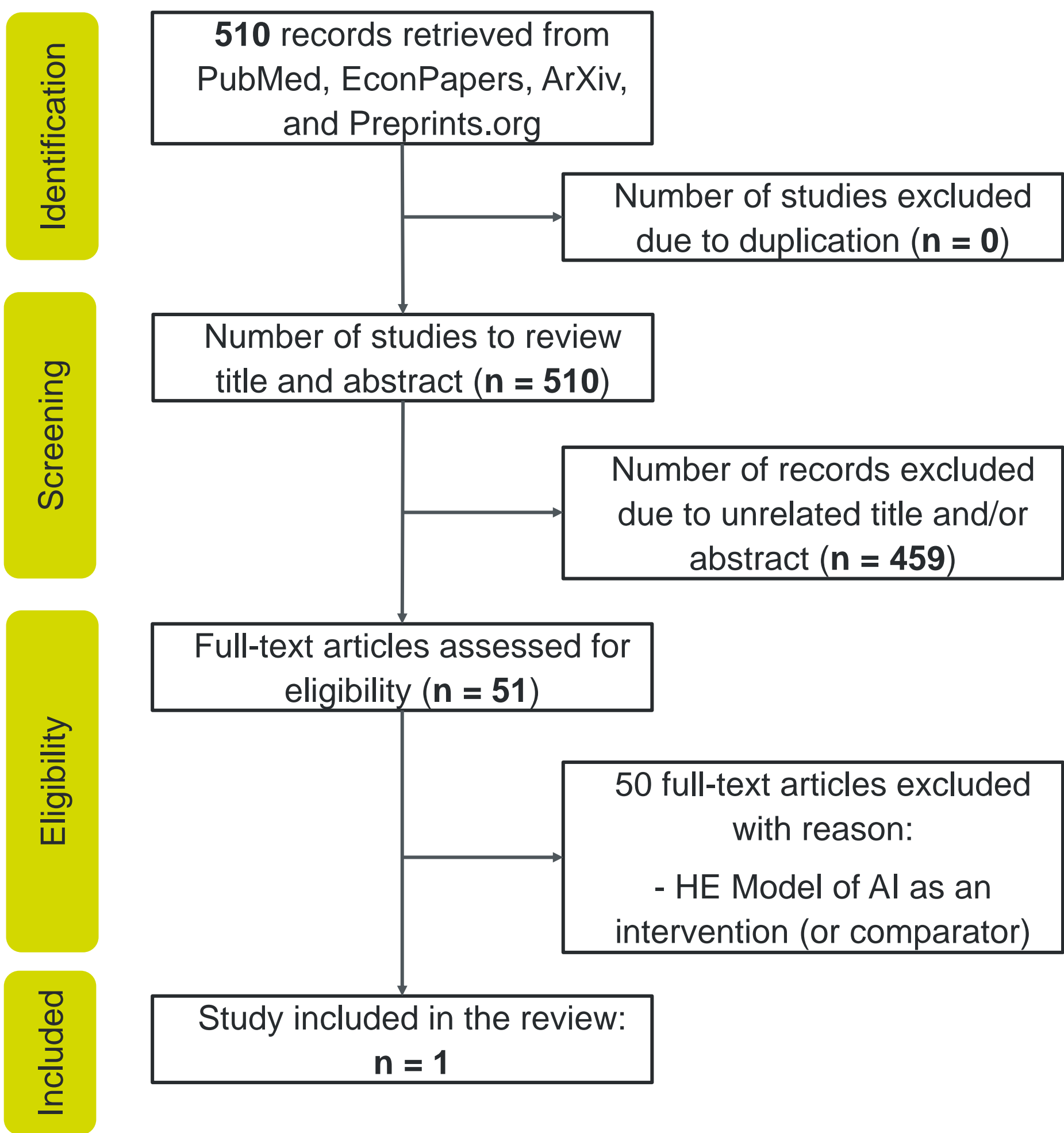
Figure 2. Composition of studies at screening
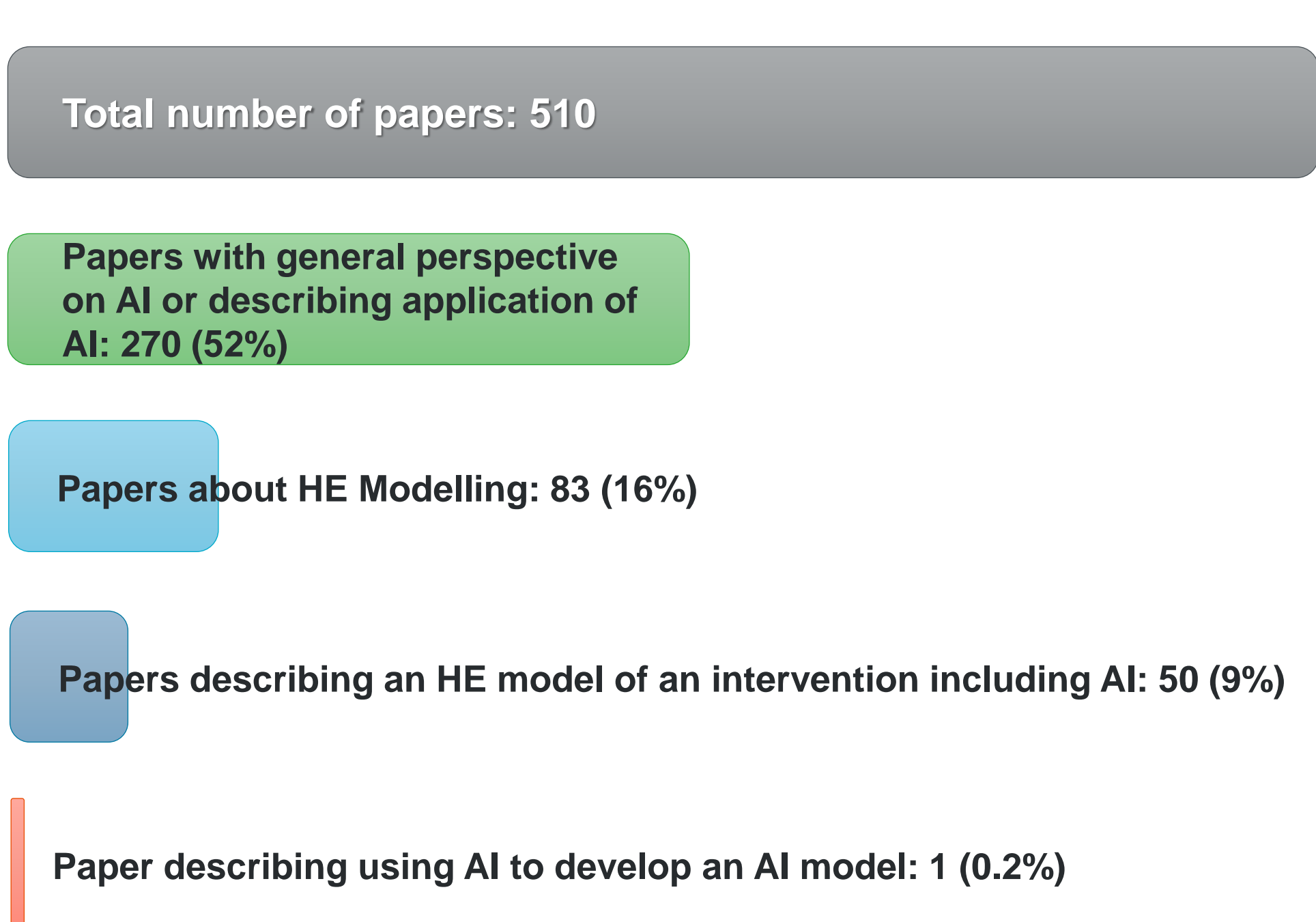
Further Research, Outlook and Conclusions

There is a high need for further research and exploration in bringing AI tools, including Language Model Models (LLMs), into the introduction of health economics modelling. The complexity of health economics modelling requires consideration of various aspects such as coding the health economic model, optimising existing code, and determining the appropriate model structure (e.g., Markov, partial survival, discrete event simulation) and the inclusion of a statement of what was the ground of truth.

There is limited reporting of AI utilisation in creating health economics models in peer-reviewed and pre-print literature, with only one recent paper available. The use of AI in health economics modelling is described, but the description is not transparent, and it remains a black box, hindering its further application and understanding. One of the intentions to introduce AI tools in modelling is to overcome the current situation of "black box models".

The feasibility of replicating oncology cost-effectiveness models in R with extensive prompt engineering is demonstrated in this paper. However, as mentioned in another study [10], reproducibility and internal and external validation, are crucial for a more robust and safer adoption of AI in health economics modelling.

A contrast is drawn between the extensive literature on AI applications in healthcare solutions and general software coding and the overly positive literature assessing the potential of AI in health economics modelling. So far, there is no proof of an advantage/improvement in the quality of HEMs using AI tools. This is also discussed in [11, 12].

In conclusion, it is recommended to explore the practical use of AI beyond LLMs for building health economics models while remaining vigilant about potential biases and other issues associated with AI utilisation. Further research is necessary to fully integrate AI tools into the development process of health economics modelling and address the existing limitations.

REFERENCES

[1] OpenAI, GPT-4 Technical Report. 2023. (doi: 10.48550/arXiv.2303.08774)

[2] Leavitt, N. and M. DeKoven. AI and HEOR: A Joining of Forces. 2021; Available from: https://jheor.org/post/1107-ai-and-heor-a-joining-of-forces

[3] Mancill, C. Embracing the Future of AI in Health Economics and Outcomes Research. 2024; Available from: https://www.linkedin.com/pulse/embracing-future-ai-health-economics-outcomes-research-chris-mancill-bspve/

[4] Malcolm, B., et al. Embracing the AI Revolution: Exploring the Promising Future of Generative AI in HEOR. in ISPOR Europe 2023. Copenhagen, Denmark.

[5] Hendrix, N., et al., Assessing the Economic Value of Clinical Artificial Intelligence: Challenges and Opportunities. Value in Health, 2022. 25(3): p. 331-339.

[6] Watters, C. and M.K. Lemanski, Universal skepticism of ChatGPT: a review of early literature on chat generative pre-trained transformer. Front Big Data, 2023. 6: p. 1224976.

[7] Sallam, M., ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. Healthcare (Basel), 2023. 11(6).

[8] Hamilton, M.P., et al., A prototype software framework for transparent, reusable and updatable computational health economic models. arXiv preprint arXiv:2310.14138, 2023.

[9] Reason, T., et al., Artificial Intelligence to Automate Health Economic Modelling: A Case Study to Evaluate the Potential Application of Large Language Models. PharmacoEcon Open, 2024. 8(2): p. 191-203.

[10] Kolasa, K., et al., Systematic reviews of machine learning in healthcare: a literature review. Expert Rev Pharmacoecon Outcomes Res, 2024. 24(1): p. 63-115.

[11] Diaz-Flores, E., et al. Evolution of Artificial Intelligence-Powered Technologies in Biomedical Research and Healthcare. In: Beutel, S., Lenk, F. (eds) Smart Biolabs of the Future. Advances in Biochemical Engineering/Biotechnology, vol 182. Springer, Cham. 2022 (doi:10.1007/10_2021_189)

[12] Sharma, S. et al.: Augmenting Health Economic Evidence Analysis: The Synergistic Role of AI in Deductive, Abductive Reasoning, and Insight Generation. 2024. Value in Health, accepted for publication

www.parexel.com

parexel®