Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/236968/

Version: Published Version

OP22 Artificial-Intelligence-Powered Prompts For Simplified Cost-Effectiveness And Net-Benefit Analysis In Health Technology Assessment

Bruno Barros (brunombarros7@gmail.com),

Carlos Alberto Magliano, Quenia Cristina Dias Morais, Marcelo Correia and Bernardo Tura

Introduction: Health technology assessment (HTA) often involves complex cost-effectiveness analyses (CEA) that can be challenging for non-experts. This study demonstrated the use of artificial-intelligence (AI)-powered prompts to simplify CEA processes. By integrating tools for generating efficiency frontiers and net-benefit analyses, this approach enables stakeholders, such as non-modeling HTA specialists or decision-makers, to understand therapeutic scenarios and make informed adjustments to analyses.

Methods: The prompts were created using ChatGPT-4o and tested for usability and reproducibility in ChatGPT-4o and the free version 3.5. The prompts were designed to automate key steps in economic analysis, including calculating net monetary and health benefits, performing cost-effectiveness analyses applying dominance and extended dominance concepts, and generating the efficient frontier plot. Ten HTA experts with no modeling experience evaluated the prompts using predefined scenarios with hypothetical datasets. The results of both versions of ChatGPT were compared to the expected results. The usability and accuracy of the prompts were assessed during the evaluation.

Results: ChatGPT-4o achieved 100 percent accuracy in calculating net health and monetary benefits (NHB and NMB) and correctly applied the concepts of dominance and extended dominance in 83 percent of cases. It calculated the incremental cost-effectiveness ratio (ICER) for non-dominated therapies in 50 percent of situations and successfully generated the efficiency frontier graphic in half of cases. In contrast, ChatGPT 3.5 achieved 50 percent accuracy for NHB and NMB calculations, only 17 percent for applying dominance and extended dominance concepts, and failed to calculate ICERs or generate the efficiency frontier plot as expected.

Conclusions: AI-powered prompts simplified cost-effectiveness analyses by enabling non-technical stakeholders to create and adjust efficiency frontiers and benefit analyses. ChatGPT-4o demonstrated improved reliability. Limitations in ChatGPT 3.5, particularly with ICER calculation and graphical outputs, indicated the need to adapt prompts to specific AI tools. Future developments may increase the robustness and usability of these types of prompts across platforms.

OP23 Developing A Personalized Decision Aid Incorporating A Discrete Choice Experiment: A Case Study In Ulcerative Colitis

Nyantara Wickramasekera (n.wickramasekera@sheffield.ac.uk)

Introduction: Choosing the optimal ulcerative colitis treatment is complex, given the range of medical and surgical options with varying side effects and effectiveness. Decision aids can improve patient choices, but current tools lack personalization. To address this, we developed a personalized decision tool using a discrete choice experiment (DCE) to help patients make informed decisions about medical or surgical treatments.

Methods: An online DCE survey was developed containing competing treatment profiles described using all important aspects of the treatment (effectiveness, side effects, family planning). Patients (n=300) with ulcerative colitis were asked to consider the benefits and disadvantages of each treatment profile and select the treatment that they would choose. The DCE data were analyzed using mixed logit and latent class models. The model results were integrated into an online decision aid using a Shiny application.

Results: R Shiny was successfully used to enable the real-time personalization of DCE results. The developed decision aid contained two aspects of personalization. First, attribute importance scores showed the treatment characteristics that mattered most to patients based on their DCE choices. Second, a "best-match" treatment that aligned with their preferences was provided from uptake rate calculations. User testing of the developed decision aid is ongoing. However, initial feedback from patients has been positive.

Conclusions: A key challenge in developing personalized decision aids is providing real-time, tailored recommendations based on individual preferences. This study demonstrated the feasibility of integrating DCE methods into personalized decision aids for ulcerative colitis. By tailoring treatment recommendations to individual patient preferences, this tool has the potential to empower patients, reduce decisional conflict, and enhance shared decision-making between patients and clinicians.