Modelling uncertain heterogeneity for decision analytic models: an early exploration

Chengyang Gao[1], Gianluca Baio[1], and Nathan Green[1]

[1] University College London

Background

- Health economic evaluations are essential for assessing the comparative costs and benefits of new health technologies;
- In health economic evaluations, it is well established that ignoring the heterogeneity of treatment effects (HTE) would lead to 'heterogeneity bias'(1, 2).
- Nevertheless, trials are rarely powered to detect HTE. The gold standard to explore HTE is a network meta-regression (3).
- To add to the uncertainty, efficacy trials only include a selective subset of the target population. Model-based extrapolation is required for HTE in the entire target population.
- Consequently, the analyses of HTE at early stages are often deemded unreliable and discouraged.

Objective

To explore whether modelling HTE is warranted for population-level decision-making when the underlying evidence is still uncertain.

The case for exploring uncertain HTE

- Exploring uncertain HTE, even at early stages, might help better represent the uncertainty of treatment benefits in the target population, as the uncertainty in the HTE can be propagate to the decision model;
- While knowledge of the HTE is scarece at early stages of drug development, the existing knowledge of the shape of the outcome surface can be formulated into structured priors to help regularise the extrapolation;
- Examples of structured priors include random walk priors which impose smoothness in the outcome surface, or multinomial logistic priors which restrict the monotoncity

Methods

General Overview

We conducted a simulation study where the analysis of a randomised controlled trial is used to inform the decision-modelling with a three-state state transition model(STM). We compare different modelling approaches with or without accounting for HTE on their ability to predict the population-level incremental net benefits (INMB) and inform decision making.

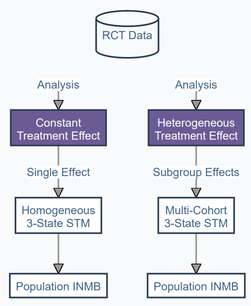


Figure 1: Flow diagram showing the simulation study setup

Trial data generation

We consider the following simplified outcome model where a binary outcome is of interest:

$$\text{logit}(p_i) = f(age_i) + g(age_i) * T_i$$

where p_i is the individual-level outcome probability. The function $f(\cdot)$ determines the shape of the control surface while $g(\cdot)$ determines the shape of the treatment surface. We consider in total 12 scenarios

We also consider *limited* and *extended* sampling scenarios where trial participants are sampled from a limited or extended part of the covariate distribution.

Outcome model of interest

Under a Bayesian framework, we consider the following outcome models for the analysis of the trial data:

- Unadjusted model
- Covariate-adjusted model
- Generalised linear model with treatment-covariate interaction
- Spline model with random walk priors (Unrestricted splines)
- Spline model with multinomial logistic priors (Monotonic splines)

Evaluation metrics:

- Decision-informing metrics: probability of cost-effectiveness at varying willingness-to-pay thresholds in comparison to the ground truth decision;
- Predictive metrics: posterior predictive distribution of mean INMB in comparisons to the true INMB

Flexible modeling of heterogeneous treatment effects with structured priors can prevent overconfidence in decision-making, even with limited evidence

When the undelrying treatment effects are decresaing over the covariate range, ignoring heterogeneity, or without structured priors, the downstream decision modelling would overestimate the probability of cost-effectiveness. This might not be mitigated by extending the sampling range.

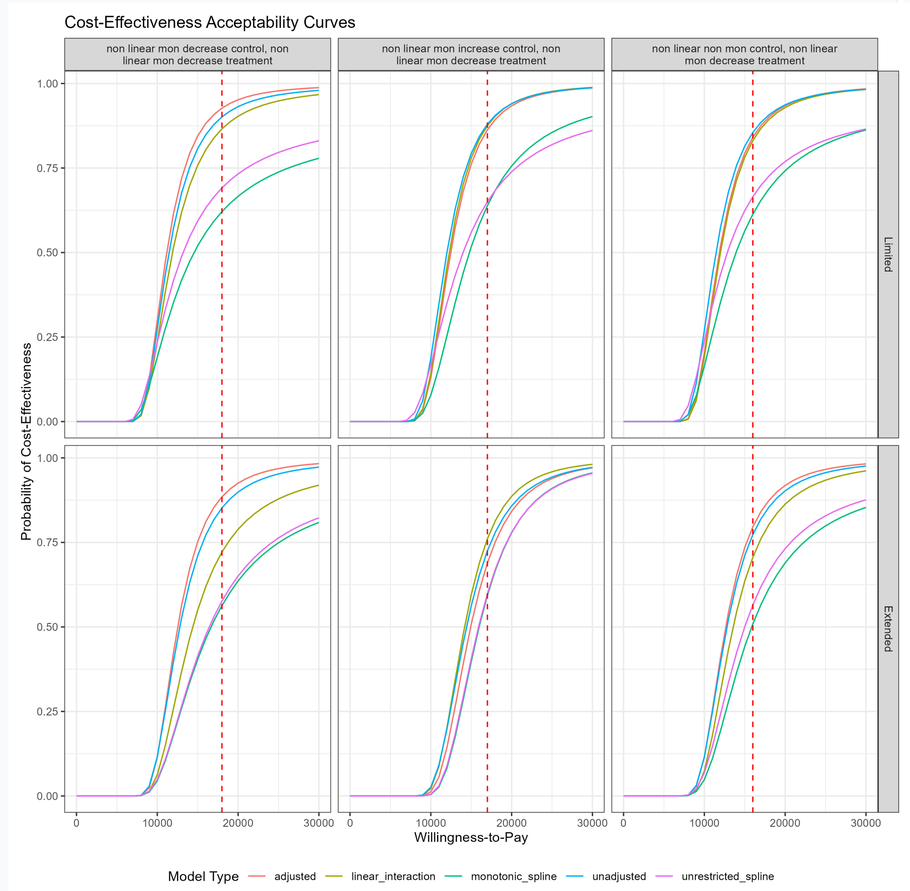


Figure 2: Cost-effectiveness acceptability curve under different modeling approaches averaged over 2000 simulations. The red dotted line indicates the willingness-to-pay threshold where the net benefits first exceed zero. Limited: only sampling from half of the covariate range; Extended: sampling from the full covariate range.

Cost-Effectiveness Acceptability Curves Across Scenarios

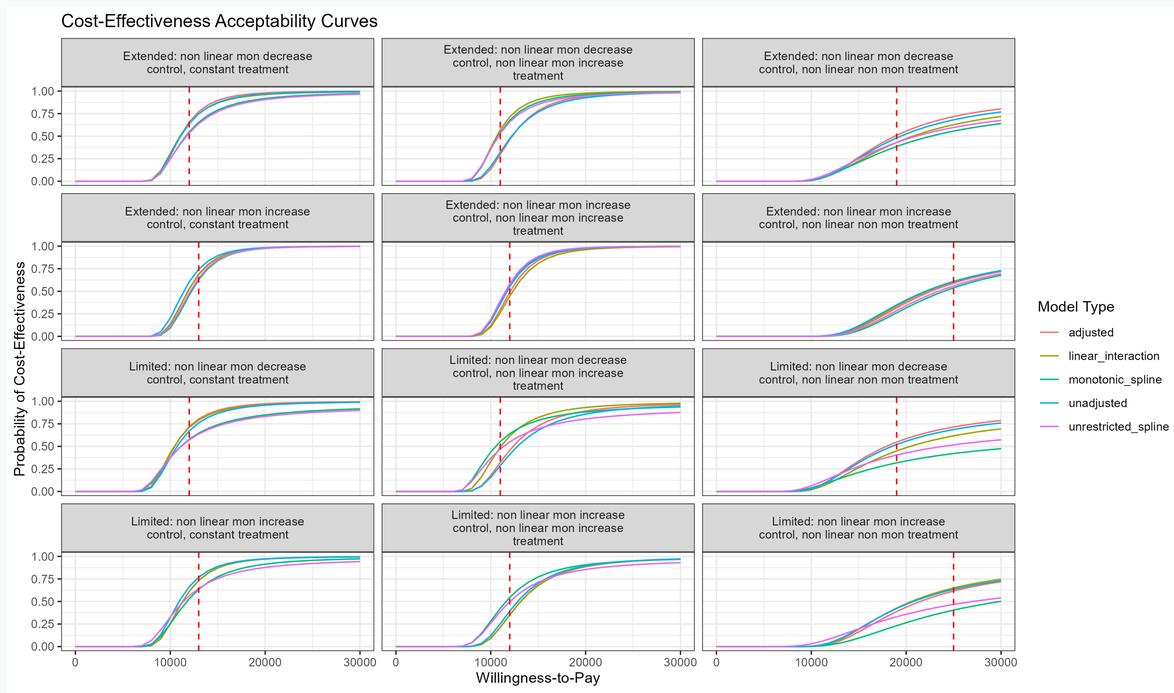
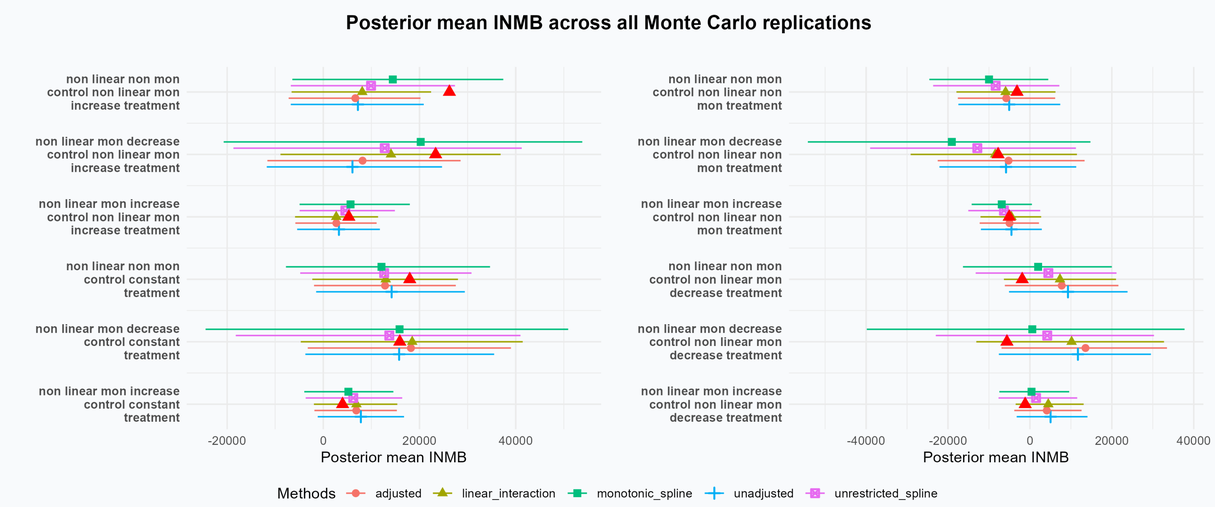


Figure 3: The remaining scenarios of the cost-effectiveness acceptability curve under different modelling approaches. Limited: individuals are sampled from a limited part of the covariate range; Extended: individuals are sampled from the full covariate range.

- The two spline models consistently assign a lower probability of cost-effectiveness across all scenarios. The predicted probabilities of cost-effectiveness based on the spline models also increase at a slower rate compared to the other outcome modelling approaches.
- The predictions from the spline models also appear to be more calibrated in most of the scenarios, assigning a closer to 50% probability of cost-effectiveness when the true INMB just becomes positive. Nevertheless, they appear to be slightly underconfident when the true treatment effect surface is non-monotonic.
- As the need for extrapolation gets reduced due to sampling from the full covariate range, there seems to be a convergence trend in the CEAC curves across all modelling approaches.
- When the true treatment effect surface is increasing over the covariate range, models ignoring HTE actually tend to be underconfident.

Posterior Predictive Distribution



Distribution of the posterior predictive mean of incremental net benefits under different modeling approaches, when individuals are sampled from a limited part of the covariate range. The red triangle indicates the true incremental net benefit.

- In terms of the predictive performance, none of the outcome modelling approaches stand out, with largely overlapping uncertainty intervals.
- while in some scenarios, the overall mean can be close to the true value, the substantial variability of posterior predictive mean across Monte Carlo replications raises questions about whether sample-based analysis alone can inform real-world decision-making, where we have to conditional on the sample at hand.

Conclusions

- Modelling HTE, even under limited supporting evidence, can mitigate over-confidence in decision-making.
- Modellers should adopt flexible modelling approaches with careful considerations of regularisation and extrapolations.
- Sample-based analysis alone has limited value in directly informing real-world decision-making, in which case communicating the uncertainty is far more important than emphasizing the magnitude of point predictions

References

1. Zaric GS. The impact of ignoring population heterogeneity when markov models are used in cost-effectiveness analysis. Medical Decision Making. Sage Publications; 2003; 23(5):379–86.
2. Kuntz KM, Goldie SJ. Assessing the sensitivity of decision-analytic results to unobserved markers of risk: Defining the effects of heterogeneity bias. Medical decision making. Sage Publications Sage CA: Thousand Oaks, CA; 2002; 22(3):218–27.
3. Dias S, Welton NJ, Sutton AJ, Caldwell DM, Lu G, Ades A. Evidence synthesis for decision making 4: Inconsistency in networks of evidence based on randomized controlled trials. Medical Decision Making. Sage Publications Sage CA: Los Angeles, CA; 2013; 33(5):641–56.

Contact Information
chengyang.gao.15@ucl.ac.uk