

A Practical Method for Determining the Optimal Allocation Between Computation Time and Accuracy in Probabilistic Patient-Level Microsimulation Models

Kirk Szafranski MSc¹, Tim Disher PhD², Stephen Brown, PhD³ • ¹EVERSANA, Burlington, ON, Canada; ²EVERSANA, Halifax, NS, Canada

Background and Objectives

- The use of probabilistic modeling increasingly informs decision making for Health Technology Assessment (HTA) agencies; for example, HTA bodies in Canada now require base cases to be probabilistic.¹
- Many HTA bodies require cost effectiveness models be programmed in Excel or other high-level software; for example, HTAs in both Canada and Australia require permission to submit models that are not programmed in Excel.^{1,2}
- Microsimulation models are a form of cost-effectiveness modeling that simulates single individuals through disease states, generating summary cohort results by combining the results of thousands or hundreds of thousands individual patient simulations.
- Probabilistic microsimulation requires re-running the initial set of patients (n) per run by a number of probabilistic Monte Carlo iterations (N).
- The computational burden of combining this approach can be immense. For example, if 10,000 simulations are generally needed for both n and N to converge, in total excel would need to run a hundred million simulations, requiring days of computation
- The objective of this research is to attempt to define a minimal number of n and N for a probabilistic microsimulation model to converge, to minimize computation time

Methods

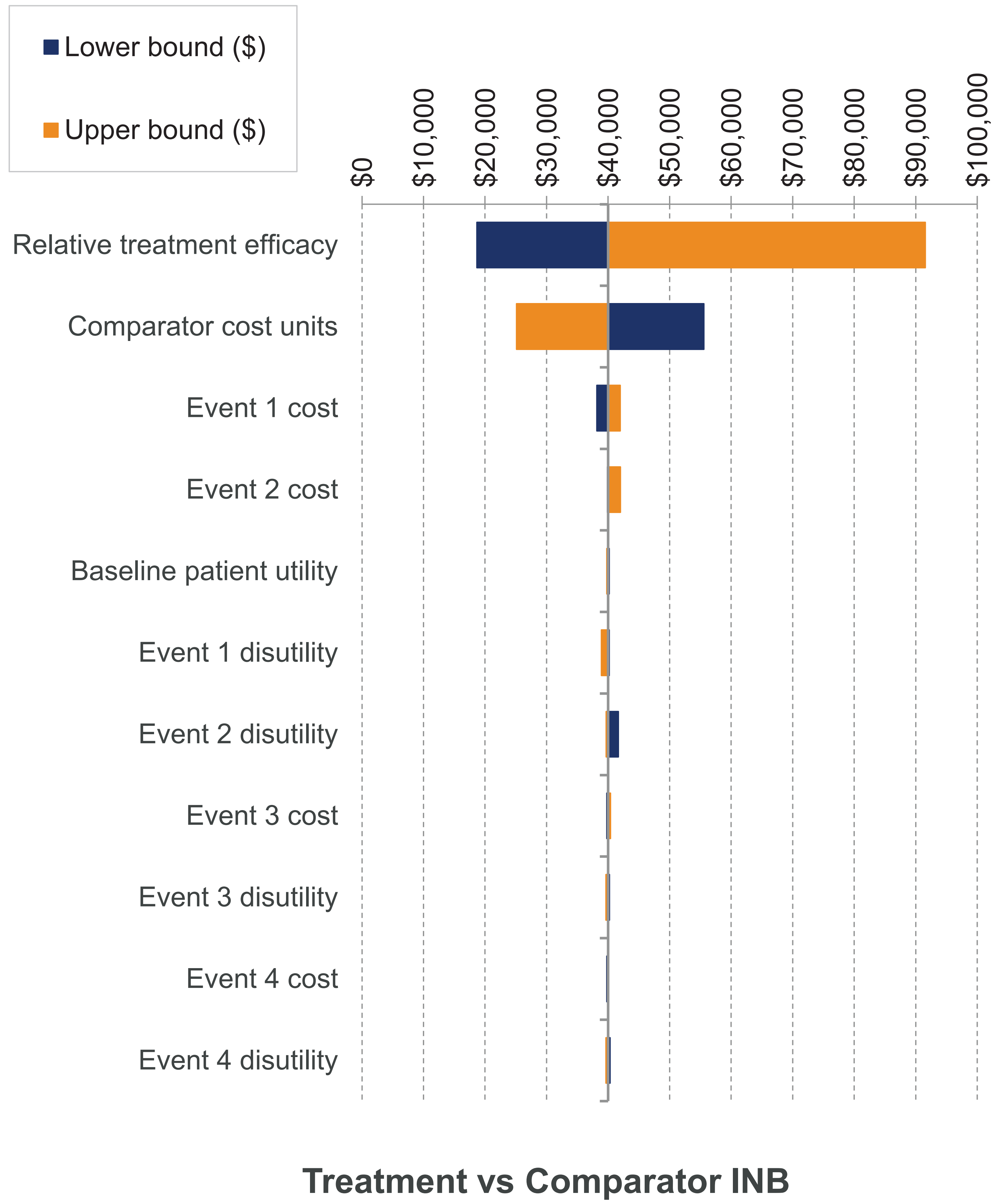
- Work done by O'Hagan et al. (2005)^{3,4} has defined methods to determine the efficient allocation of n and N which result in a combined computation burden M (M=n*N) such that a given sampling precision c_2 is reached for a probabilistic microsimulation model.
- M can be calculated as $8k/c_2^2$ where k is a ratio of the mean patient-level and probabilistic run-level variance of the incremental net benefit (INB) of a therapy.
- Determining k (**Figure 1**) requires creating initial runs of combinations of model inputs ± 1.5 * their standard deviation – to evaluate all combinations requires 3^d initial runs, where d is the number of parameters run in the probabilistic microsimulation.
- Most probabilistic models submitted to HTAs have tens to hundreds of inputs whose uncertainty HTA bodies wish to assess. For example, the model tested in this approach has 79 inputs, which would result in 3^{79} or over 10^{37} initial runs to determine k. Sampling at random has been proposed but may put undue weight on variables whose uncertainty does not impact results
- Instead we propose to identify the most impactful variables using one-way sensitivity analyses
- this approach was tested using one to three of the most impactful variables in a diabetes microsimulation using n = 100,000 patients (which was sufficient to give less than 0.1% uncertainty in each microsimulation run).

Figure 1: Steps to determine sampling burden M for a given precision c_2

- Step 1**
Run a one-way sensitivity analysis (OWSA) and select the two or three most impactful variables (d = 2 or 3)
- Step 2**
Create a 3^d matrix of all possible combinations of each selected variable ± 1.5 standard deviation (SD).
- Step 3**
Run the microsimulation model for each with a sufficient n and record each run's mean INB and patient-level variance in INB
- Step 4**
Average the patient-level variance in the INB over all 3^d runs (τ^2)
- Step 5**
Determine the variance in the mean of each of the 3^d runs (σ_i). The probabilistic run level variance σ^2 can be calculated using:
$$\sigma^2 = \left(\sigma_i^2 - \frac{\tau^2}{n} \right) \left(\frac{2}{3} \right)^d$$
- Step 6**
Calculate k as the ratio of τ^2 and σ^2
$$k = \frac{\tau^2}{\sigma^2}$$
- Step 7**
Determine M, n and N for a given precision c_2 using the following equations:
$$M = \frac{8k}{c_2^2} \qquad n = 1 + k$$
$$N = \frac{M}{n}$$

Results

Figure 2: OWSA results



Discussion

Selecting the top variables is a practical approach to determine optimal n and N for probabilistic microsimulations. This approach allows the most impactful variables to determine the threshold for convergence in probabilistic microsimulations. However, examining the run-times in Table 2 in our example probabilistic microsimulation model demonstrates that the use of Excel to run these analyses may be impractical. Alternative approaches such as meta-models may better enable practical run-times for probabilistic microsimulation models programmed in Excel.⁵

Table 1: Calculation of k using one to three variables

Variables tested	k calculation
Relative treatment efficacy	12,370
Relative treatment efficacy, comparator unit costs	18,555
Relative treatment efficacy, comparator unit costs, Event 1 costs	18,579

Summary

- K calculations were very similar when using two or three variables (**Figure 2** and **Table 1**), so we proceeded with using the two most impactful variables, relative treatment efficacy and number of cost units of the comparator.
- Our OWSA (**Figure 1**) showed that these variables had a disproportionate impact on the INB when compared with subsequent variables in the OWSA. OWSA that reveal a more even distribution in impact between variables may need to stop at a practical d even if k doesn't converge.
- We estimated run-time for our d=2 result given a range of precision (c_2) desired. Run-time estimates are presented in **Table 2**.

Table 2: Probabilistic iterations and run-times for different precision levels

Precision (c_2)	Patients simulated (n)	Probabilistic iterations (N)	Run-time in Excel
0.2	18,556	200	3.9 days
0.1	18,556	800	15.5 days
0.05	18,556	3,200	61.9 days
0.01	18,556	80,000	1,546 days

Abbreviations HTA = Health Technology Assessment; INB = incremental net benefit; OWSA = one-way sensitivity analysis; SD = standard deviation.

References 1. CADTH (2023). Procedures for CADTH Reimbursement Reviews - April 2023. 2. PBAC (2016). Guidelines for preparing a submission to the Pharmaceutical Benefits Advisory Committee Version 5.0. 3. O'Hagan A, Stevenson M, Madan J (2005). Monte Carlo probabilistic sensitivity analysis for patient level simulation models. [White Paper]. 4. O'Hagan A, Stevenson M, Madan J (2007) Monte Carlo probabilistic sensitivity analysis for patient level simulation models: efficient estimation of mean and variance using ANOVA. *Health Econ* 16 (10): 1009-1023. 5. Jalal H, Goldhaber-Fiebert JD, Kuntz KM (2015) Computing Expected Value of Partial Sample Information from Probabilistic Sensitivity Analysis Using Linear Regression Metamodeling. *Med Decis Making* 35 (5): 584-595.

Presented at the 2023 ISPOR US Conference
Boston MA – May 7 to 10, 2023

