General-Purpose Methods for Simulating Survival Data for Expected Value of Sample Information Calculations

Running head: Simulating Survival Data for EVSI Calculations

18 January, 2023

Abstract

Background Expected value of sample information (EVSI) quantifies the expected value to a decision maker of reducing uncertainty by collecting additional data. EVSI calculations require simulating plausible datasets, typically achieved by evaluating quantile functions at random uniform numbers using standard inverse transform sampling (ITS). This is straightforward when closed-form expressions for the quantile function are available, such as for standard parametric survival models, but these are often unavailable when assuming treatment effect waning and for flexible survival models. In these circumstances, the standard ITS method could be implemented by numerically evaluating the quantile functions at each iteration in a probabilistic analysis, but this greatly increases the computational burden. Thus, our study aims to develop general-purpose methods that standardize and reduce the computational burden of the EVSI data-simulation step for survival data. **Methods** We developed a discrete sampling method and an interpolated ITS method for simulating survival data from a probabilistic sample of survival probabilities over discrete time units. We compared the general-purpose and standard ITS methods using an illustrative partitioned survival model with and without adjustment for treatment effect waning. **Results** The discrete sampling and interpolated ITS methods agree closely with the standard ITS method, with the added benefit of a greatly reduced computational cost in the scenario with adjustment for treatment effect waning. **Conclusions** We present general-purpose methods for simulating survival data from a probabilistic sample of survival probabilities that greatly reduce the computational burden of the EVSI data-simulation step when we assume treatment effect waning or use flexible survival models. The implementation of our data-simulation methods is identical across all possible survival models and can easily be automated from standard probabilistic decision analyses.

Introduction

Expected Value of Sample Information (EVSI) quantifies the expected value to a decision maker of reducing uncertainty through a given data collection exercise, such as a randomized clinical trial[1]. Methods for computing the EVSI for collecting survival data (i.e. time-to-event data) when there is uncertainty about the choice of survival model have recently been developed by Vervaart et al[2]. These methods require, in common with other EVSI methods, the simulation of plausible study datasets that reflect the study design proposed for collecting future data and the time-to-event distribution of individuals included in such a study[3]. This is typically achieved by evaluating quantile functions at random uniform numbers using standard inverse transform sampling (ITS). The standard ITS method is straightforward to implement when closed-form expressions for the quantile function are available, such as for standard parametric survival models, but these are often not available when assuming treatment effect waning and for flexible survival models.

Sufficient evidence on time-to-event outcomes, such as overall survival and time to progression, is crucial for accurately determining the long-term effects of new treatments.[4] Yet, health technology assessments often have to rely on immature survival data obtained from trials at an early stage, especially for new cancer treatments[5]. This can partly be explained by the introduction of accelerated licensing schemes for new pharmaceuticals by regulatory bodies such as the European Medicines Agency[6,7] and the U.S. Food and Drug Administration[8]. Immature survival data require a high degree of extrapolation, which have led to the introduction of flexible survival models such as response-based landmark models, mixture cure models, relative survival models, and model-averaging approaches.[9,10] Nevertheless, more complex models do not necessarily result in plausible extrapolations, and therefore extrapolations are often supplemented with assumptions about disease progression and treatment mechanisms. For example, the National Institute for Health and Care Excellence (NICE) recommends that waning of treatment effects is considered in technology appraisals[11], for instance by assuming no more treatment benefit beyond a chosen time point[12], or by assuming that the treatment effect diminishes over the long term[13]. This is typically implemented in cost-effectiveness models by adjusting the predicted hazards, thereby altering the survival probabilities generated by parametric survival models. This poses a challenge for the EVSI data-simulation step for survival data, as closed-form expressions for the quantile function are often unavailable for custom distributions that incorporate assumptions about treatment effect waning and for flexible survival models. In these circumstances, the standard ITS method for simulating survival data could be implemented by numerically evaluating the quantile distributions at each iteration in a probabilistic analysis, but this can greatly increase the computational burden.

In this article, we address the problem of computing EVSI when we assume treatment effect waning or use flexible survival models, by developing general-purpose methods that standardize and reduce the computational burden of the EVSI data-generation step for survival data. We develop a discrete sampling method and an interpolated ITS method for simulating survival data from a probabilistic sample of discrete survival curve probabilities. The discrete sampling method samples time cycles using the discrete survival curve probabilities and sets the event times to the half-cycle times. The interpolated ITS method extends this to continuous time by initially sampling random uniform numbers between 0 and 1, and then interpolating the discrete survival curve probabilities using cubic splines at the sampled numbers and recording the interpolated cycle times. We demonstrate in an illustrative case study that, when the general-purpose data-simulation methods are combined with a recently proposed nonparametric EVSI method[2,14], EVSI computations for survival data can be automated from standard probabilistic decision analyses, irrespective of the assumed data-generating process.

The article is structured as follows. In the second section, we first describe the standard ITS method and then introduce the general-purpose methods for simulating survival data. In the third section, we compare the standard ITS method and the general-purpose data-simulation methods based on an illustrative partitioned survival model for scenarios with and without adjustment for treatment effect waning. In the final section, we conclude with a short discussion.

Method

Decision problem

In health technology assessment, cost-effectiveness models are widely used to compare alternative health technologies in terms of expected costs V and health benefits Q relative to a cost-effectiveness threshold λ.[15] There is usually a choice between a small number of $\mathcal{D}$ decision options, indexed $d = 1, \dots \mathcal{D}$, such as a new treatment and standard care. The cost-effectiveness of decision option d can be expressed in terms of net health benefit, $\mathrm{NB}_d = Q_d - V_d/\lambda$, or net monetary benefit, $\mathrm{NB}_d = Q_d\lambda - V_d$. A cost-effectiveness model, which we denote as $\mathrm{NB}_d(\boldsymbol{\theta})$, predicts the net benefit of decision option d, given a vector of model input parameters $\boldsymbol{\theta}$, such as probabilities, costs and health-related quality of life weights. We represent our current knowledge about the cost-effectiveness model parameters in the joint probability distribution $p(\boldsymbol{\theta})$. The joint distribution $p(\boldsymbol{\theta})$ is typically defined using a combination of statistical models fitted to individual patient data, external data sources, clinical expert opinion and assumptions about biological plausibility, and there may also be dependency between elements of $\boldsymbol{\theta}$.

EVSI for a new study

The expected value of the optimal decision given current information is the value of the decision option that maximizes expected net benefit and can therefore be considered cost-effective,

$$\max_d \mathbb{E}_{\boldsymbol{\theta}}\{\text{NB}_d(\boldsymbol{\theta})\}, \tag{1}$$

A new study would provide data $\mathbf{X}$ relating to $\boldsymbol{\theta}$, which could for example be a randomized controlled trial that collects overall survival data for a new treatment and standard care. Once the study data have been collected, we update the joint distribution of the cost-effectiveness model parameters with the new data using Bayes theorem, giving a posterior distribution $p(\boldsymbol{\theta}|\mathbf{X})$.

The expected value of the optimal decision made after collecting $\mathbf{X}$ is

$$\max_d \mathbb{E}_{\boldsymbol{\theta}|\mathbf{X}}\{\text{NB}_d(\boldsymbol{\theta})\}. \tag{2}$$

The value of collecting $\mathbf{X}$ is derived from its potential to help a decision maker avoid recommending a treatment that reduces net benefit, as we may learn that the decision option that is considered cost-effective given current information turns out to be cost-ineffective given new information. However, as we have not yet collected $\mathbf{X}$, we must take the expectation over the distribution of all possible datasets, $p(\mathbf{X})$. This requires simulating plausible datasets from the distribution of the new data, $\mathbf{x} \sim p(\mathbf{X})$. To simulate $\mathbf{x}$, we usually need to specify a *parametric* data-generating distribution, $p(\mathbf{X}|\boldsymbol{\theta})$, which depends on the type of data that will be generated by the proposed study and how it will be used to update the model parameters.[3] We can simulate datasets from $p(\mathbf{X})$ by first sampling from the joint distribution of the model parameters, $\boldsymbol{\theta}^* \sim p(\boldsymbol{\theta})$, and then sampling from the parametric data-generating distribution given the parameter sample, $\mathbf{x}^* \sim p(\mathbf{X}|\boldsymbol{\theta}^*)$. This gives us a pair of samples $\{\mathbf{x}^*, \boldsymbol{\theta}^*\}$ from the joint distribution $p(\mathbf{X}, \boldsymbol{\theta})$, and therefore the samples $\mathbf{x}^*$ are draws from the marginal distribution $p(\mathbf{X})$.

The expected value of the decision made with additional sample information is given by,

$$\mathbb{E}_{\mathbf{X}}[\max_d \mathbb{E}_{\boldsymbol{\theta}|\mathbf{X}}\{\text{NB}_d(\boldsymbol{\theta})\}]. \tag{3}$$

The EVSI[16] measures the expected value of reducing uncertainty about the optimal decision by collecting $\mathbf{X}$, which is defined as the difference between equation (3) and equation (1),

$$\text{EVSI} = \mathbb{E}_{\mathbf{X}}[\max_d \mathbb{E}_{\boldsymbol{\theta}|\mathbf{X}}\{\text{NB}_d(\boldsymbol{\theta})\}] - \max_d \mathbb{E}_{\boldsymbol{\theta}}\{\text{NB}_d(\boldsymbol{\theta})\}. \tag{4}$$

In the next section, we describe the EVSI estimation procedure for collecting time-to-event data.

Computing EVSI for time-to-event data

Time-to-event data

Time-to-event data, such as time to disease progression and time to death, are frequently collected in the context of clinical trials. A special feature of time-to-event data is censoring, which occurs when the follow-up time is not long enough to observe the event of interest for all individuals, or when individuals are lost to follow-up.[17] A single time-to-event dataset $\mathbf{x}$, collected between time 0 and t_{obs}, consists of $i = 1, \ldots, n$ survival times x_i and censoring indicators δ_i, $\mathbf{x} = \{x_1, \ldots, x_n, \delta_1, \ldots, \delta_n\}$, where $\delta_i = 1$ when x_i is an observed event, and $\delta_i = 0$ when x_i is a censored observation.

To predict outcomes over the long term, censored time-to-event data usually need to be extrapolated beyond the observed follow-up period using a parametric survival model[18]. Parametric models are commonly specified using either the survivor function, $S(t, \boldsymbol{\theta})$, or hazard function, $h(t, \boldsymbol{\theta})$, with time denoted as t.

The survivor function $S(t)$ defines the probability of survival up to time t, given by

$$S(t, \boldsymbol{\theta}) = P(T > t) = 1 - F(t, \boldsymbol{\theta}),\ 0 < t < \inf, \tag{5}$$

where $F(t, \boldsymbol{\theta})$ is the cumulative distribution function.

The hazard function $h(t, \boldsymbol{\theta})$ defines the instantaneous event rate at time t conditional on survival up to time t,

$$h(t, \boldsymbol{\theta}) = \lim_{\delta t \to 0} \left\{ \frac{P(t \leq T < t + \delta t \mid T \geq t)}{\delta t} \right\} = \frac{f(t, \boldsymbol{\theta})}{S(t, \boldsymbol{\theta})}, \tag{6}$$

where $f(t, \boldsymbol{\theta})$ is the probability density function.

Most cost-effectiveness models are in discrete-time and therefore evaluate $S(t, \boldsymbol{\theta})$ at fixed time intervals based on a set amount of time called model cycles, thereby generating a vector of survival probabilities over discrete time units for each sampled value for $\boldsymbol{\theta}$, from which the distribution of patients across the health states over time can be constructed. The expected costs, life years and quality-adjusted life years of a new technology can be estimated by integrating the distribution of patients across the health states and the health state values, which reflect the costs and health-related quality of life weights associated with each health state.[15]

Computing the expected net benefits given current information

We can compute the expected net benefits given current information in a probabilistic analysis (PA) using Monte Carlo simulation. This involves sampling $k = 1, \ldots, K$ values, $\boldsymbol{\theta}^{(k)}$, from the distribution of the model parameters, $p(\boldsymbol{\theta})$, and then evaluating the survivor functions, $S(t_c, \boldsymbol{\theta}^{(k)})$, given model cycle $c = 0, \ldots, C$. This results in K vectors of survival probabilities, $\mathbf{s}^{(k)} = \{s_0^{(k)}, \ldots, s_C^{(k)}\}$, where $s_c^{(k)} = S(t_c, \boldsymbol{\theta}^{(k)})$. We assume $\mathbf{s} \subset \boldsymbol{\theta}$, and compute the net benefits conditional on the survival curves and other model parameters, $\text{NB}_d(\boldsymbol{\theta}^{(k)})$, for each d. We compute the expected net benefits given current information by averaging over the K net benefits for each d.

Table 1 illustrates a PA sample where K vectors of survival probabilities $\mathbf{s}^{(k)}$ have been generated for a single decision option d.

Table 1: Probabilistic analysis sample with K sampled vectors of survival probabilities $\mathbf{s}^{(k)}$ for a single decision option.

Model cycle	$\mathbf{s}^{(1)}$	$\mathbf{s}^{(2)}$	$\mathbf{s}^{(3)}$	$\ldots$	$\mathbf{s}^{(K)}$
0	1.00	1.00	1.00	$\ldots$	1.00
1	0.85	0.79	0.88	$\ldots$	0.92
2	0.76	0.70	0.79	$\ldots$	0.88
3	0.69	0.63	0.72	$\ldots$	0.79
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
C	0.00	0.00	0.00	$\ldots$	0.00
$\sum_{c=0}^{C} s_c^{(k)}$	14.49	10.82	12.11	$\ldots$	18.79

Simulating Time-to-Event Data

Standard inverse transform sampling method

To compute the expected net benefits given new time-to-event data $\mathbf{x}$, we must first sample plausible datasets from the distribution of the new data conditional on the sampled parameter values, $\mathbf{x}^{(k)} \sim p(\mathbf{X}|\boldsymbol{\theta}^{(k)})$. We can simulate random survival times from a parametric survival distribution using standard functions, such as the `rweibull` function in `R`[19]. These standard functions use inverse transform sampling (ITS), which is a method for generating random values from an arbitrary distribution. Let F be a continuous cumulative distribution function (CDF) with inverse CDF (i.e. quantile function) denoted by F^{-1}. The intuition behind the ITS method is that if we define $T = F^{-1}(U)$, where $U \sim \text{Unif}[0,1]$, then T is distributed according to F, i.e. $F(T)$. Since the survivor function is defined as $S(t,\boldsymbol{\theta}) = 1 - F(t,\boldsymbol{\theta})$, we can generate survival times by sampling values $u_i,\ i = 1,\ldots,n$, from a uniform distribution on the interval $[0,1]$ and plugging these into the inverse survivor function, $S^{-1}(u_i,\boldsymbol{\theta})$. The inverse survivor function can be derived from the inverse CDF, $S^{-1}(1-u_i,\boldsymbol{\theta}) = F^{-1}(u_i,\boldsymbol{\theta})$, and if $U \sim \text{Unif}[0,1]$, it follows that $1-U$ is also $\sim \text{Unif}[0,1]$. We then censor the survival times at the new follow-up time t_{new}.

The standard ITS scheme for simulating time-to-event data is given in Box 1.

Box 1: Standard inverse transform sampling scheme for simulating time-to-event data

```
for $k = 1,\ldots,K$ parameter samples $\boldsymbol{\theta}^{(k)}$ do
    Sample a vector of values $\mathbf{u}^{(k)} = \{u_1^{(k)},\ldots,u_n^{(k)}\}$, with each element drawn from Unif$[0,1]$
    Generate a dataset of $n$ survival times $\mathbf{x}^{(k)} = \{x_1^{(k)},\ldots,x_n^{(k)}\}$ by plugging $\mathbf{u}^{(k)}$ into $S^{-1}(\mathbf{u}^{(k)},\boldsymbol{\theta}^{(k)})$
    Censor each element of $\mathbf{x}^{(k)}$ at the new follow-up time $t_{new}$
end
```

Sampling from a Weibull distribution, such as by using the `rweibull` function in `R`, can result in any value between 0 and infinity. If we want to ensure that the sampled survival times do not exceed a biologically plausible time horizon t_h, which would typically be equal to the time horizon used in the decision model, then we need to sample $\mathbf{u}_i^{(k)}$ on the interval $[p_1^{(k)},1]$, where $p_1^{(k)}$ is the survivor function evaluated at t_h, i.e. $S\left(t_h,\boldsymbol{\theta}^{(k)}\right)$. If, however, the chance of simulating implausible survival times that exceed t_h is substantial, the analyst should consider using a more realistic survival model, in particular models that incorporate background mortality and other external information[9,20]. However, this evidence may not be available and is exactly the evidence that is under consideration to collect in the value of information analysis.

The algorithm in Box 1 could, in theory, be used for any survival model for which we can define a hazard function, $h(t,\boldsymbol{\theta})$. The hazard could be integrated to produce a cumulative hazard function, $H(t,\boldsymbol{\theta})$, from which we can derive the survivor function, $S(t,\boldsymbol{\theta}) = \exp[-H(t,\boldsymbol{\theta})]$. The survivor function could then be inverted to produce the quantile function, $S^{-1}(\cdot,\boldsymbol{\theta})$, and survival times could be simulated by evaluating $S^{-1}(\cdot,\boldsymbol{\theta})$ at a random uniform sample as described in Box 1.

Analytic solutions to the integrals and function inverses, such as implemented in the `rweibull` function, may not be available for flexible survival models, such as relative survival models, spline models, mixture cure models and response-based landmark models[9,10], and for custom distributions, including hazard functions that incorporate assumptions about treatment effect waning. In these circumstances, the integrals and function inverses could be evaluated numerically, for example by using the `integrate()` and `uniroot()` functions in base R. The `flexsurv` and `msm` packages on CRAN also have a function `qgeneric()` designed to invert a generic CDF. These numerical methods, however, greatly increase the computational burden of the EVSI computations, as the integration and inversion steps need to be repeated for each of $k = 1,\ldots,K$ simulations. In Appendix A, we provide a simple example and step-by-step implementation in `R` of the standard ITS method based on analytic and numerical solutions. The example illustrates that even for a simple exponential model, numerically evaluating the integrals and function inverses results in over a 10,000-fold increase in computational time (Table A1).

In the next section, we will introduce general-purpose methods for simulating survival data that can be standardized from standard probabilistic decision analyses and greatly reduce the computational burden of the standard ITS method when closed-form expressions for the quantile function are unavailable.

Simulating time-to-event data from a vector of survival probabilities over discrete time units

Interpolated inverse transform sampling method

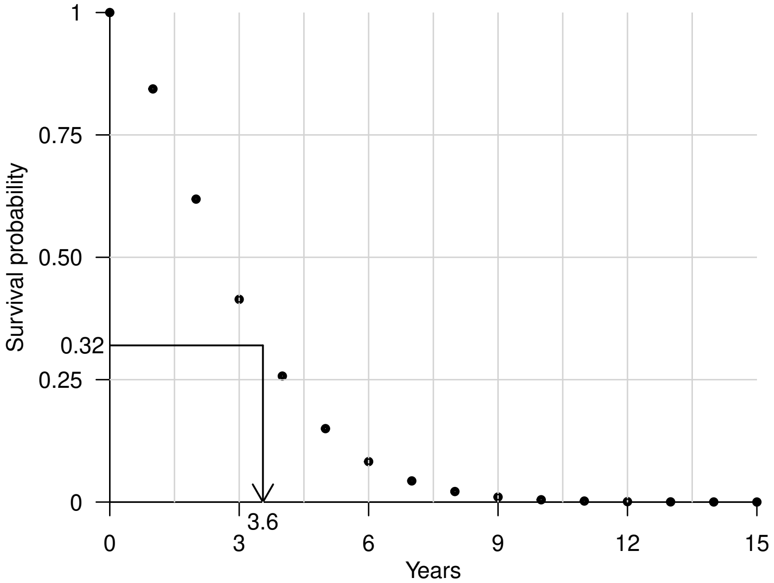


Figure 1: Illustration of the interpolated inverse transform sampling method for simulating time-to-event data. A survival time of 3.6 years has been simulated by first sampling a value of 0.32 from a uniform distribution between 0 and 1, and then interpolating the survival probabilities over discrete time units using monotone cubic splines at 0.32 and recording the interpolated cycle time of 3.6 years.

We can also use the ITS method to generate $\mathbf{x}$ from a vector of survival probabilities over discrete time units, as illustrated in Figure 1. Our PA sample consists of K vectors of survival probabilities, $\mathbf{s}^{(k)} = \{s_0^{(k)}, \ldots, s_C^{(k)}\}$, given model cycle $c = 0, \ldots, C$. We can approximate values from the quantile function by interpolating the vectors of survival probabilities, $S^{-1}(\cdot, \mathbf{s}^{(k)})$. We can achieve this by fitting monotone cubic splines, a type of piecewise polynomial interpolation, to each consecutive set of survival probabilities and cycle times, $\{(s_0^{(k)}, s_1^{(k)}), (t_0, t_1)\}, \ldots, \{(s_{C-1}^{(k)}, s_C^{(k)}), (t_{C-1}, t_C)\}$. We can then sample survival times from the approximate quantile function $S^{-1}(\cdot, \mathbf{s}^{(k)})$ using the ITS method.

The interpolated ITS sampling scheme for simulating time-to-event data from a vector of survival probabilities over discrete time units is given in Box 2.

Box 2: Interpolated inverse transform sampling scheme for simulating time-to-event data from a vector of survival probabilities over discrete time units

```
for $k = 1, \ldots, K$ vectors of survival probabilities $\mathbf{s}^{(k)} = \{s_0^{(k)}, \ldots, s_C^{(k)}\}$ do
    Sample a vector of values $\mathbf{u}^{(k)} = \{u_1^{(k)}, \ldots, u_n^{(k)}\}$, with each element drawn from Unif$[s_C^{(k)}, 1]$
    Generate a dataset of $n$ survival times $\mathbf{x}^{(k)} = \{x_1^{(k)}, \ldots, x_n^{(k)}\}$ by interpolating $\mathbf{s}^{(k)}$ at $\mathbf{u}^{(k)}$ using
      monotone cubic splines and recording the interpolated cycle times
    Censor each element of $\mathbf{x}^{(k)}$ at the new follow-up time $t_{new}$
end
```

Discrete sampling method

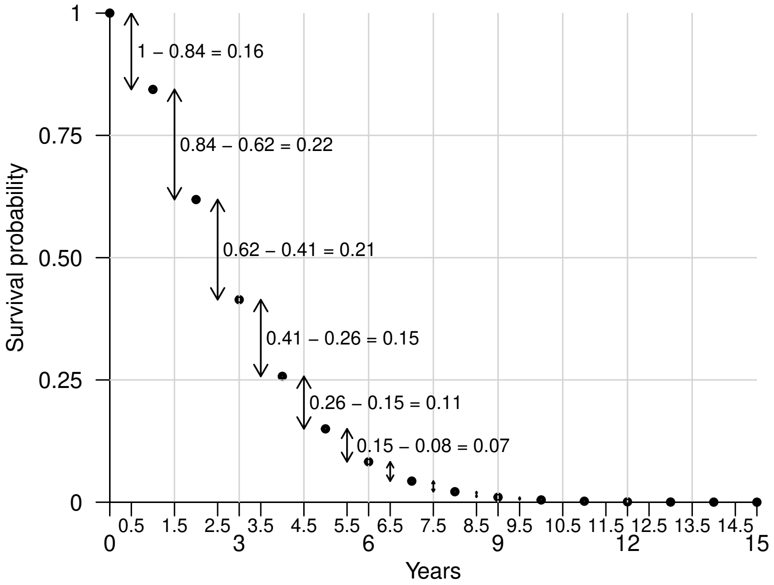


Figure 2: Illustration of the discrete sampling method for simulating time-to-event data. Survival times can be simulated by sampling from the half-cycle times on the x-axis with probability derived from the survival probabilities over discrete time units as indicated by the arrows.

An alternative approach for simulating $\mathbf{x}$ from a vector of survival probabilities over discrete time units $\mathbf{s}^{(k)}$ is by sampling discrete 'bins' of cycle times with probability equal to the cumulative density of each bin, which is illustrated in Figure 2. We set the value of each bin equal to the half-cycle time $t_{c+0.5}$, $c = 0, \ldots, C-1$, i.e. the midpoint in between each consecutive set of cycle times. We can estimate the cumulative density of each bin from the survival probabilities, $(s_c^{(k)} - s_{c+1}^{(k)})$. We can then generate survival times by sampling $\{t_{0.5}, \ldots, t_{C-0.5}\}$ with probability $\{(s_0^{(k)} - s_1^{(k)}), \ldots, (s_{C-1}^{(k)} - s_C^{(k)})\}$.

The discrete sampling scheme for simulating time-to-event data from a vector of survival probabilities over discrete time units is given in Box 3.

Box 3: Discrete sampling scheme for simulating time-to-event data from a vector of survival probabilities over discrete time units

```
for $k = 1, \ldots, K$ vectors of survival probabilities $\mathbf{s}^{(k)} = \{s_0^{(k)}, \ldots, s_C^{(k)}\}$ do
    Generate a dataset of $n$ survival times $\mathbf{x}^{(k)} = \{x_1^{(k)}, \ldots, x_n^{(k)}\}$ by sampling $n$ values from the
      half-cycle times, $\{t_{0.5}, \ldots, t_{C-0.5}\}$, with probability $\{(s_0^{(k)} - s_1^{(k)}), \ldots, (s_{C-1}^{(k)} - s_C^{(k)})\}$
    Censor each element of $\mathbf{x}^{(k)}$ at the new follow-up time $t_{new}$
end
```

Appendix A describes a step-by-step implementation in R of the interpolated ITS and discrete sampling methods, and a comparison of their computational efficiency with the standard ITS method based on analytic and numerical solutions.

Computing the expected net benefits given new data

We could compute the expected net benefits given the $k = 1, \ldots, K$ sampled datasets for each decision option, $\mathbf{x}_d^{(k)}$, using a nested Monte Carlo scheme, but this can be very computationally demanding as it requires sampling a large number of values from the posterior distribution of the model parameters conditional on each simulated dataset, $p(\boldsymbol{\theta}|\mathbf{x}_d^{(k)})$. A number of efficient approximation methods have therefore been developed in recent years that reduce the computational burden of the nested Monte Carlo procedure to EVSI[21]. One of these efficient EVSI methods is a nonparametric regression-based method developed by Strong et al. (2015)[14] which does not require a parametric distribution for the data. The regression-based method for computing EVSI relies on estimating the functional relationship between the posterior expected net benefits and the simulated datasets, thereby avoiding the need to sample from the posterior distributions $p(\boldsymbol{\theta}|\mathbf{x}_d^{(k)})$, as is required for the nested Monte Carlo scheme.

In the regression-based approach, we only require the vectors of $k = 1, \ldots, K$ prior net benefits $\text{NB}_d(\boldsymbol{\theta}^{(k)})$ for each decision option d that we generated in the PA and the corresponding datasets $\mathbf{x}_d^{(k)}$ that we have simulated using either of the sampling schemes above. The observed net benefits $\text{NB}_d(\boldsymbol{\theta}^{(k)})$ can be expressed as a sum of the conditional expectation of the net benefit given the data, $\mathbb{E}_{\boldsymbol{\theta}|\mathbf{x}_d^{(k)}}\{\text{NB}_d(\boldsymbol{\theta})\}$, which we require to estimate the EVSI (Equation (4)), and a mean-zero error term, $\varepsilon^{(k)}$,

$$\text{NB}_d(\boldsymbol{\theta}^{(k)}) = \mathbb{E}_{\boldsymbol{\theta}|\mathbf{x}_d^{(k)}}\{\text{NB}_d(\boldsymbol{\theta})\} + \varepsilon^{(k)}. \tag{7}$$

Strong and others (2015)[14] explain that the conditional expectation $\mathbb{E}_{\boldsymbol{\theta}|\mathbf{x}_d^{(k)}}\{\text{NB}_d(\boldsymbol{\theta})\}$ can be thought of as an unknown function of $\mathbf{x}_d^{(k)}$. We denote this function $g(\mathbf{x}_d^{(k)})$ and substitute this into Equation (7), giving

$$\text{NB}_d(\boldsymbol{\theta}^{(k)}) = g(\mathbf{x}_d^{(k)}) + \varepsilon^{(k)}. \tag{8}$$

We then summarize $\mathbf{x}_d^{(k)}$ using a low-dimensional summary statistic for each d, $T(\mathbf{x}_d^{(k)})$,

$$\text{NB}_d(\boldsymbol{\theta}^{(k)}) = g\{T(\mathbf{x}_d^{(k)})\} + \varepsilon^{(k)}. \tag{9}$$

A convenient choice for $T(\mathbf{x}_d^{(k)})$ is the number of observed events $e_d^{(k)}$ and the total time at risk $y_d^{(k)}$ for each simulated dataset $\mathbf{x}_d^{(k)}$, i.e. $T(\mathbf{x}_d^{(k)}) = \{e_d^{(k)}, y_d^{(k)}\}$, which has been shown to give good results for various survival models.[2]

We can estimate the posterior net benefits by regressing the prior net benefits, $\text{NB}_d(\boldsymbol{\theta}^{(k)})$, on the summary statistic $T(\mathbf{x}_d^{(k)})$,

$$\text{NB}_d(\boldsymbol{\theta}^{(k)}) = g\{T(\mathbf{x}_d^{(k)})\} + \varepsilon, \tag{10}$$

where g_d is a function of the data $\mathbf{x}_d^{(k)}$ for each d, and ε is an error term with zero mean. We can achieve this by fitting a generalized additive model (GAM), which is a flexible nonparametric regression model, to each d and extracting the regression model fitted values $\hat{g}_d^{(k)}$, which are estimates of the posterior net benefits.

The GAM-based EVSI estimate is given by

$$\text{EVSI} \simeq \frac{1}{K}\sum_{k=1}^{K}\max_d \hat{g}_d^{(k)} - \max_d \frac{1}{K}\sum_{k=1}^{K}\hat{g}_d^{(k)}. \tag{11}$$

EVSI for an ongoing study

When a trial is ongoing at the point of decision making, there could be value in reducing uncertainty by collecting additional data from the ongoing trial before making an adoption decision. This is especially common for cost-effectiveness analyses of new cancer drugs, which increasingly rely on immature data obtained from trials in an early stage[5].

We denote the new data collected between the observed follow-up time t_{obs} and future point t_{new} as $\tilde{\mathbf{x}} = \{\tilde{x}_1, \ldots, \tilde{x}_{\tilde{n}}, \tilde{\delta}_1, \ldots, \tilde{\delta}_{\tilde{n}}\}$, where $\tilde{n}$ is the number of study participants at risk at t_{obs}. As we have not yet collected the new data, we need to simulate plausible new data sets conditional on the current data, $\tilde{\mathbf{x}} \sim (\tilde{\mathbf{X}}|\mathbf{X})$. The value of extending an existing trial's follow-up from t_{obs} to future point t_{new} is given by[2]

$$\text{EVSI(ongoing study)} = \mathbb{E}_{\tilde{\mathbf{X}}|\mathbf{X}}[\max_d \mathbb{E}_{\boldsymbol{\theta}|\mathbf{X},\tilde{\mathbf{X}}}\{\text{NB}(d,\boldsymbol{\theta})\}] - \max_d \mathbb{E}_{\boldsymbol{\theta}|\mathbf{X}}\{\text{NB}(d,\boldsymbol{\theta})\}, \tag{12}$$

where the first term is the expected value of a decision based on the joint posterior distribution of $\boldsymbol{\theta}$ given both new data, $\tilde{\mathbf{X}}$, collected between t_{obs} and t_{new}, and current data, $\mathbf{X}$, collected between time zero and t_{obs}, which is computed by averaging over the posterior net benefits of the decision option that maximizes this quantity conditional on both new data and current data. The second term is the expected value of a decision based on the joint distribution of $\boldsymbol{\theta}$ given current data collected up until t_{obs}.

Events beyond t_{obs} are conditional on survival up to t_{obs}. Therefore, to simulate future survival times for $\tilde{n}$ patients at risk at t_{obs}, we need to sample from a conditional distribution that is left-truncated at t_{obs}. For the standard ITS scheme in Box 1, this requires evaluating the survivor function at t_{obs}, $p_2^{(k)} = S\left(t_{obs}, \boldsymbol{\theta}^{(k)}\right)$, and then sampling random uniform numbers, $\mathbf{u}^{(k)}$, on the interval $[p_1^{(k)}, p_2^{(k)}]$, where $p_1^{(k)}$ is the survivor function evaluated at the model time horizon, $S\left(t_h, \boldsymbol{\theta}^{(k)}\right)$. For the interpolated ITS scheme in Box 2, we can find $p_2^{(k)}$ by interpolating $\mathbf{s}^{(k)}$ at t_{obs}, and then sampling $\mathbf{u}^{(k)}$ on the interval $[s_C^{(k)}, p_2^{(k)}]$. For the discrete sampling scheme in Box 3, we need to sample from the subset of half-cycle times that are larger than t_{obs}, i.e. $\{t_{c+0.5}, \ldots, t_{C-0.5}\}$ for $t_{c+0.5} > t_{obs}$.

Model-averaged EVSI

Uncertainty about the choice of survival model is often a key driver of decision uncertainty, particularly when data are immature[22]. If we are uncertain about choosing from a set of competing survival models for extrapolating study data over the long-term, $\mathcal{M} = M_r,\ r = 1, \ldots, R$, then we could account for this in the EVSI computations[2] by using model-averaging.[23–25] Before we collect new data $\mathbf{X}$, our beliefs about the plausibility of each model is represented by the prior model probabilities $P(M_r)$. These could for example be derived from the Akaike's Information Criterion (AIC)[26] or other measures of model fit and parsimony. In the PA, we then sample a survival model $M_r^{(k)}$ with probability $P(M_r)$ before sampling $\boldsymbol{\theta}_r^{(k)}$ from the distribution of the parameters of the sampled survival model, $p(\boldsymbol{\theta}_r, M_r^{(k)})$. Since this only changes the values for $\mathbf{s}^{(k)}$, the discrete sampling scheme in Box 2 and interpolated ITS scheme in Box 3, as well as the GAM-based EVSI estimation following equation (11), are identical to the single-model case.

Synthetic case study

Decision problem and model definition

To demonstrate the application of our methods, we developed a simple yet realistic synthetic case study based on a partitioned survival model (PSM)[27] comparing a new treatment ($d = 1$) with standard care ($d = 2$). The PSM uses overall survival (OS) and progression-free survival (PFS) curves to estimate the proportion of patients in 3 health states: PFS, post-progression survival (PPS) and death, given $c = 0, \ldots, 360$ monthly model cycles corresponding to an overall time horizon $t_h = 30$ years. We assumed OS and PFS follow independent Weibull distributions for each d, parameterized in terms of log-shape α and log-scale β. The Weibull model parameters for OS are $\boldsymbol{\theta}_{os1} = (\alpha_{os1}, \beta_{os1})$ for the new treatment and $\boldsymbol{\theta}_{os2} = (\alpha_{os2}, \beta_{os2})$ for standard care, and $\boldsymbol{\theta}_{pfs1} = (\alpha_{pfs1}, \beta_{pfs1})$ and $\boldsymbol{\theta}_{pfs2} = (\alpha_{pfs2}, \beta_{pfs2})$ for PFS, respectively. We estimated the Weibull model parameters using maximum likelihood from a synthetic data set containing 100 OS times and 100 PFS times for each trial arm with a maximum follow-up of 24 months. Further details about the synthetic case study data set are given in appendix B. The other model parameters are utility for PFS (U_{pfs}), utility for post-progression (U_{pps}), drug costs for the new treatment (C_{drug1}) in PFS, medical costs for the new treatment (C_{med1}), medical costs for standard care (C_{med2}), annual discount rate (r) and monetary value of 1 quality-adjusted life year (QALY), (λ). The definitions and prior distributions for the case study model parameters are given in Table 2, and net benefit functions are given in appendix C.

Table 2: Prior Parameter Distributions for the Partitioned Survival Model Parameters

Weibull survival model parameters		Mean, $\boldsymbol{\mu}$	Covariance matrix, $\boldsymbol{\Sigma}$	Distribution
Overall survival				
Log shape for new treatment Log scale for new treatment	$\begin{pmatrix}\alpha_{os1}\\ \beta_{os1}\end{pmatrix}$	$\begin{pmatrix}0.312\\ 4.089\end{pmatrix}$	$\begin{pmatrix}0.037 & -0.036\\ -0.036 & 0.056\end{pmatrix}$	Bivariate Normal($\boldsymbol{\mu}, \boldsymbol{\Sigma}$)
Log shape for standard care Log scale for standard care	$\begin{pmatrix}\alpha_{os2}\\ \beta_{os2}\end{pmatrix}$	$\begin{pmatrix}0.361\\ 3.842\end{pmatrix}$	$\begin{pmatrix}0.029 & -0.021\\ -0.021 & 0.030\end{pmatrix}$	Bivariate Normal($\boldsymbol{\mu}, \boldsymbol{\Sigma}$)
Progression-free survival				
Log shape for new treatment Log scale for new treatment	$\begin{pmatrix}\alpha_{pfs1}\\ \beta_{pfs1}\end{pmatrix}$	$\begin{pmatrix}0.161\\ 3.590\end{pmatrix}$	$\begin{pmatrix}0.019 & -0.010\\ -0.010 & 0.021\end{pmatrix}$	Bivariate Normal($\boldsymbol{\mu}, \boldsymbol{\Sigma}$)
Log shape for standard care Log scale for standard care	$\begin{pmatrix}\alpha_{pfs2}\\ \beta_{pfs2}\end{pmatrix}$	$\begin{pmatrix}0.209\\ 3.294\end{pmatrix}$	$\begin{pmatrix}0.014 & -0.004\\ -0.004 & 0.013\end{pmatrix}$	Bivariate Normal($\boldsymbol{\mu}, \boldsymbol{\Sigma}$)
Utility parameters		Mean, $\boldsymbol{\mu}$	Standard error, **SE**	Distribution
Utility for progression-free survival	U_{pfs}	0.80	0.04	Beta(80,20)
Utility for post-progression survival	U_{pps}	0.50	0.05	Beta(50,50)
Cost parameters (monthly)		Mean, $\boldsymbol{\mu}$	Standard error, **SE**	Distribution
Drug costs for new treatment	C_{drug1}	$1,200$	–	Constant
Medical costs for new treatment	C_{med1}	500	250	Gamma(4,0.008)
Medical costs for standard care	C_{med2}	500	250	Gamma(4,0.008)
Other parameters		Mean, $\boldsymbol{\mu}$	Standard error, **SE**	Distribution
Annual discount rate	r	0.035	–	Constant
Monetary value of 1 QALY	λ	$80,000$	–	Constant

Treatment stopping rule and treatment effect waning

We also considered a scenario with a 2-year treatment stopping rule, after which the drug costs for the new treatment $C_{drug1} = 0$. We assumed that, on expectation, the treatment effect on OS and PFS would wane after stopping treatment at year 2 until there was no more treatment effect by year 4. Treatment effect waning is commonly implemented in cost-effectiveness models by setting the hazards in the treatment arm and comparator arm equal at a chosen time point. This approach has a number of limitations. Firstly, this could lead to counter-intuitive results when the hazard in the comparator arm is below the hazard in the treatment arm after the treatment duration cut-off, in which case the treatment effect increases during the waning period[12]. Secondly, setting the hazards in the treatment and comparator arm equal underestimates uncertainty about independent survival endpoints. Thirdly, this approach ignores uncertainty about the start time and duration of the waning period.

To avoid these limitations, we used an alternative approach to implement treatment effect waning. For $k = 1, \ldots, K$, we first sampled values, $t_{w1}^{(k)}$ and $t_{w2}^{(k)}$, for the start time and duration of the waning period from $p(t_{w1}) \sim \text{LogNormal}(2.83, 0.69)$ and $p(t_{w2}) \sim \text{LogNormal}(2.83, 0.69)$, corresponding to mean times and standard deviations of 24 months for both t_{w1} and t_{w2}. In practical applications, $p(t_{w1})$ and $p(t_{w2})$ could be estimated from previous studies or by expert elicitation[28]. We then computed vectors of 'waning hazards' from the Weibull mean survival probabilities for the new treatment and standard care that in linear proportions increase from no waning to full waning over the time periods defined by $t_{w1}^{(k)}$ and $t_{w2}^{(k)}$, and added these to the sampled Weibull hazards for the new treatment, $\mathbf{h}_1^{(k)}$.

The waning-adjusted hazard function for OS and PFS for the new treatment is given by

$$\text{Waning-adjusted hazard } h_1^*(t) = \frac{e^{\theta_{\alpha 1}}}{e^{\theta_{\beta 1}}}\left(\frac{t}{e^{\theta_{\beta 1}}}\right)^{e^{\theta_{\alpha 1}}-1} + F_{Unif}\left(\frac{t-t_{w1}}{t_{w2}} \mid 0,1\right)\max[0, f_{spline}(t, \hat{\mathbf{h}}_2 - \hat{\mathbf{h}}_1)], \quad (13)$$

where the first term is the Weibull hazard function, $F_{Unif}(\,\cdot \mid 0,1)$ is a Uniform cumulative distribution function on the interval $[0,1]$ evaluated at $(t-t_{w1})/t_{w2}$, and $f_{spline}(t, \hat{\mathbf{h}}_2 - \hat{\mathbf{h}}_1)$ is a spline function fitted to the difference in mean survival hazards for the new treatment and standard care, $\hat{\mathbf{h}}_1$ and $\hat{\mathbf{h}}_2$, respectively, with mean survival hazards for treatment option d given by

$$\text{Mean survival hazards } \hat{\mathbf{h}}_d = -\log\left(1 - \frac{\{\bar{s}_{d,1}, \ldots, \bar{s}_{d,C}\}}{\{\bar{s}_{d,0}, \ldots, \bar{s}_{d,C-1}\}}\right), \quad (14)$$

where $\bar{s}_{d,c}$, $c = 0, \ldots, C$, are mean Weibull survival probabilities over discrete time units for each d, given by

$$\text{Mean survival probability } \bar{s}_{d,c} = \frac{1}{K}\sum_{k=1}^{K} e^{-\left(t_c/e^{\beta_d^{(k)}}\right)^{e^{\alpha_d^{(k)}}}}. \quad (15)$$

All other model assumptions and net benefit functions are as above.

The expected Weibull survival curves for the scenarios with and without adjustment for treatment effect waning are given in Figure 3.

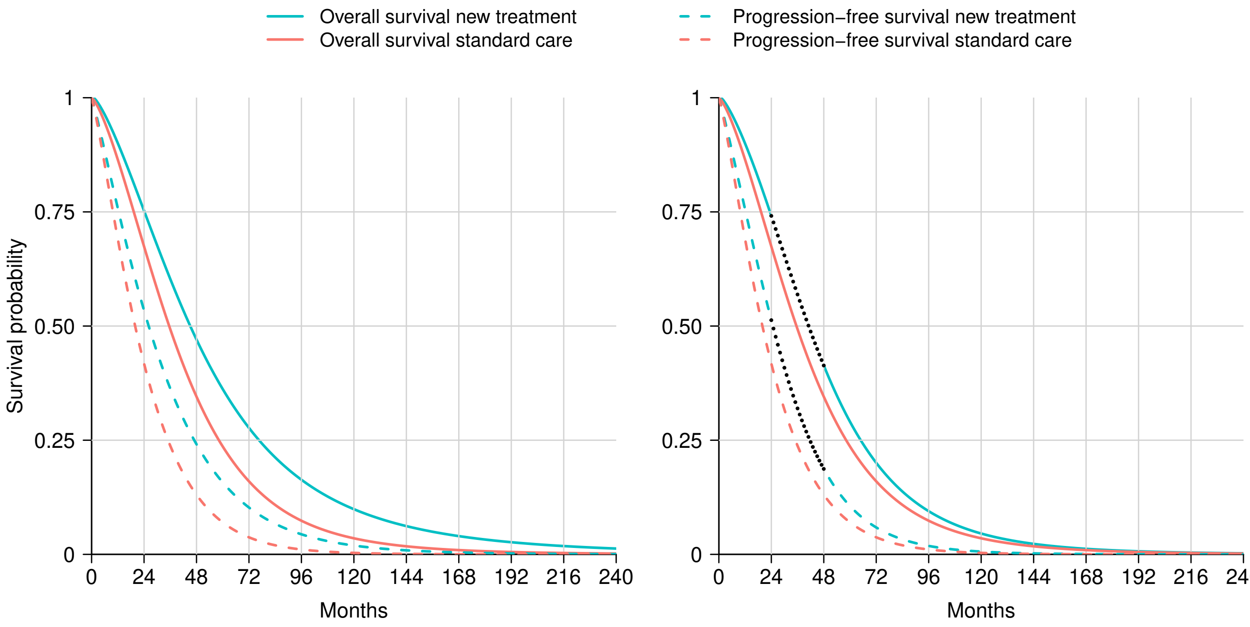


Figure 3: Expected Weibull survival curves for overall survival and progression-free survival for the new treatment and for standard care without adjustment (left) and with adjustment for treatment effect waning (right). The expected waning period is represented by the black dotted lines.

Computations

We assume we want to compute the EVSI for a new study that will collect OS and PFS data for the new treatment and for standard care. We considered a sample size of $n = 100$ study participants for each treatment arm and a study follow-up period $t_{new} = 1, 2, 3, 4, 5$ and 10 years. We first performed a PA in which we sampled $K = 2,000$ parameter values $\boldsymbol{\theta}^{(k)}$ from the distribution of the model parameters $p(\boldsymbol{\theta})$, and then evaluated the PSM to obtain the net benefits for each d, $\text{NB}_d(\boldsymbol{\theta}^{(k)})$.

Simulating overall survival and progression-free survival data

In the scenario without adjustment for treatment effect waning, we simulated OS data sets, $\mathbf{x}_{os1}^{(k)}$ for the new treatment and $\mathbf{x}_{os2}^{(k)}$ for standard care, with $n = 100$ times for each treatment arm using the standard ITS method, the interpolated ITS method and the discrete sampling method following the sampling schemes in Box 1, Box 2 and Box 3, respectively. In the scenario with adjustment for treatment effect waning, we implemented the standard ITS method by numerically integrating the waning hazard function and inverting the waning survivor function, as analytic solutions are unavailable. The implementation of the interpolated ITS method and the discrete sampling method is identical to the scenario without treatment effect waning. We summarized each simulated data set using the number of observed OS events, $e_{os1}^{(k)}$ and $e_{os2}^{(k)}$, and the total time at risk for OS, $y_{os1}^{(k)}$ and $y_{os2}^{(k)}$, i.e. $T(\mathbf{x}_{os1}^{(k)}) = \{e_{os1}^{(k)}, y_{os1}^{(k)}\}$ and $T(\mathbf{x}_{os2}^{(k)}) = \{e_{os2}^{(k)}, y_{os2}^{(k)}\}$.

We simulated PFS data sets for the new treatment and standard care ($\mathbf{x}_{pfs1}^{(k)}$ and $\mathbf{x}_{pfs2}^{(k)}$) using the same sampling schemes as for OS for the scenarios with and without adjustment for treatment effect waning. To prevent double-counting of OS data, we censored the simulated PFS times at the proposed follow-up time or at the time point at which, if at all, the PFS curve crosses and is set equal to the OS curve, whichever is soonest. We then computed the number of observed PFS events, $e_{pfs1}^{(k)}$ and $e_{pfs2}^{(k)}$, and the total time at risk for PFS, $y_{pfs1}^{(k)}$ and $y_{pfs2}^{(k)}$, for each simulated PFS data set, i.e. $T(\mathbf{x}_{pfs1}^{(k)}) = \{e_{pfs1}^{(k)}, y_{pfs1}^{(k)}\}$ and $T(\mathbf{x}_{pfs2}^{(k)}) = \{e_{pfs2}^{(k)}, y_{pfs2}^{(k)}\}$.

Computing EVSI via GAM regression

To reduce the number of regression equations[14] and improve the stability of the EVSI computations[29], we used the incremental net benefit (INB), defined as $\text{INB}(\boldsymbol{\theta}^{(k)}) = \text{NB}_1(\boldsymbol{\theta}^{(k)}) - \text{NB}_2(\boldsymbol{\theta}^{(k)})$. We estimated posterior INB by fitting a single GAM model with $\text{INB}(\boldsymbol{\theta}^{(k)})$ as the dependent variable and separate sets of summary statistics per treatment arm, $\{e_{os1}^{(k)}, y_{os1}^{(k)}, e_{pfs1}^{(k)}, y_{pfs1}^{(k)}\}$ and $\{e_{os2}^{(k)}, y_{os2}^{(k)}, e_{pfs2}^{(k)}, y_{pfs2}^{(k)}\}$, as the independent variables. We implemented the GAM regression using the `R` package mgcv[30], and specified a tensor product cubic regression spline basis for the independent variables, with a maximum basis dimension of 4 to prevent the model from estimating too many coefficients. This has syntax `gam(inb ~ te(e_os1, y_os1, e_pfs1, y_pfs1, k=4) + te(e_os2, y_os2, e_pfs2, y_pfs2, k=4)`. We extracted the GAM model fitted values, $\hat{g}^{(k)}$, which are estimates of the posterior INB, and estimated EVSI using the equation given by

$$\text{EVSI} \simeq \frac{1}{K}\sum_{k=1}^{K} \max\{0, \hat{g}^{(k)}\} - \max \frac{1}{K}\sum_{k=1}^{K}\{0, \hat{g}^{(k)}\}. \qquad (16)$$

We computed 95% intervals for the GAM estimator by sampling 2,000 values from a multivariate Normal distribution of the GAM coefficients, as described in an appendix of the article by Strong et al.[31].

Results

Figure 4 shows the EVSI values and 95% intervals without adjustment for treatment effect waning for follow-up times of 1, 2, 3, 4, 5 and 10 years. There is excellent agreement between the standard ITS method, interpolated ITS method and discrete sampling method both for OS only and OS and PFS. The EVSI reflects diminishing marginal returns for increasing follow-up durations, ranging from 2,711 to 11,311 for OS only and from 3,387 to 11,547 for OS and PFS, and converges towards the partial EVPI for the respective sets of model parameters. This indicates that the value of reducing uncertainty about PFS in addition to OS is relatively small. We did not compute the EVSI for PFS only, since PFS is a composite endpoint that is defined as time to progression or time to death, whichever is soonest, and therefore also requires the collection of OS data. The total computation times for the data-simulation procedures in the scenario without adjustment for treatment effect waning are 12 s, 32 s and 12 s for the standard ITS method, interpolated ITS method and discrete sampling method, respectively.

The EVSI estimates in the scenario with adjustment for treatment effect waning (Figure 5) are greater than in the scenario without adjustment for treatment effect waning, reflecting the added value of learning about treatment effect waning. The EVSI estimates range from 5,433 to 11,961 for OS only and from 6,064 to 12,384 for OS and PFS, almost twice as high for the 1 year follow-up period compared to the scenario without adjustment for treatment effect waning. The interpolated ITS method and discrete sampling method again agree closely with the standard ITS method, but at a greatly reduce computational cost. The computation times for the interpolated ITS and discrete sampling methods are, in fact, the same as in the scenario without adjustment for treatment effect waning, and approximately 3,600 and 10,000 times faster, respectively, than the standard ITS scheme that used numerical solutions for the integrals and function inverses.

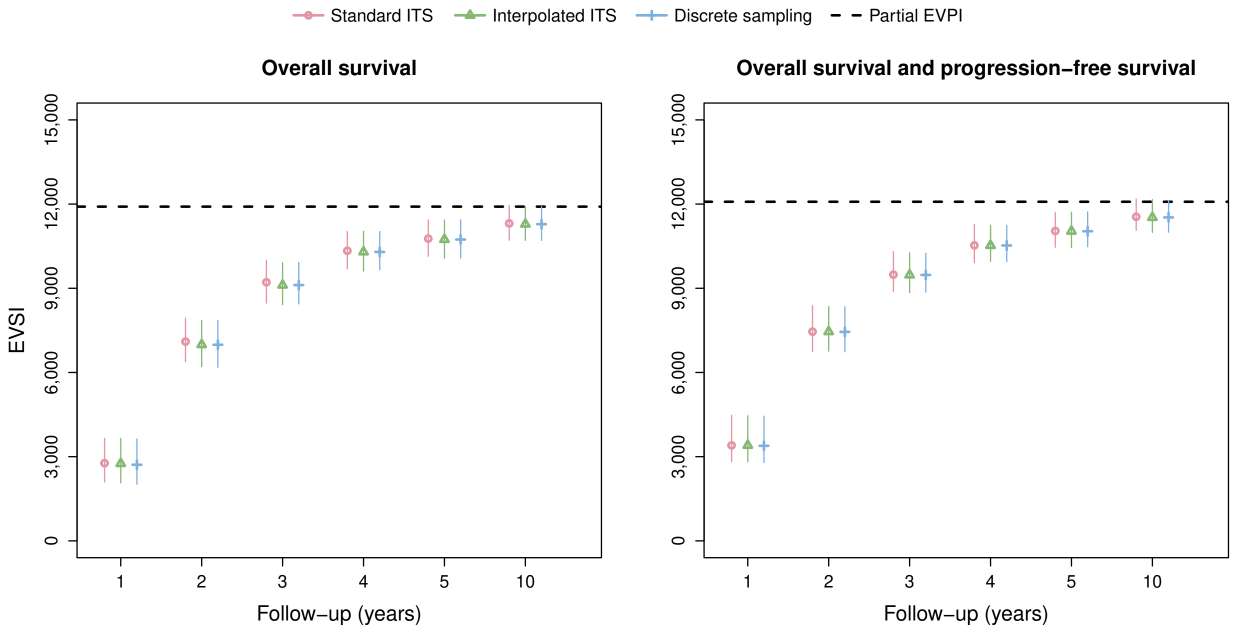


Figure 4: EVSI values for the synthetic case study without adjustment for treatment effect waning. Total computation times for the data-simulation procedures are 12 s (standard ITS), 32 s (interpolated ITS) and 12 s (discrete sampling).

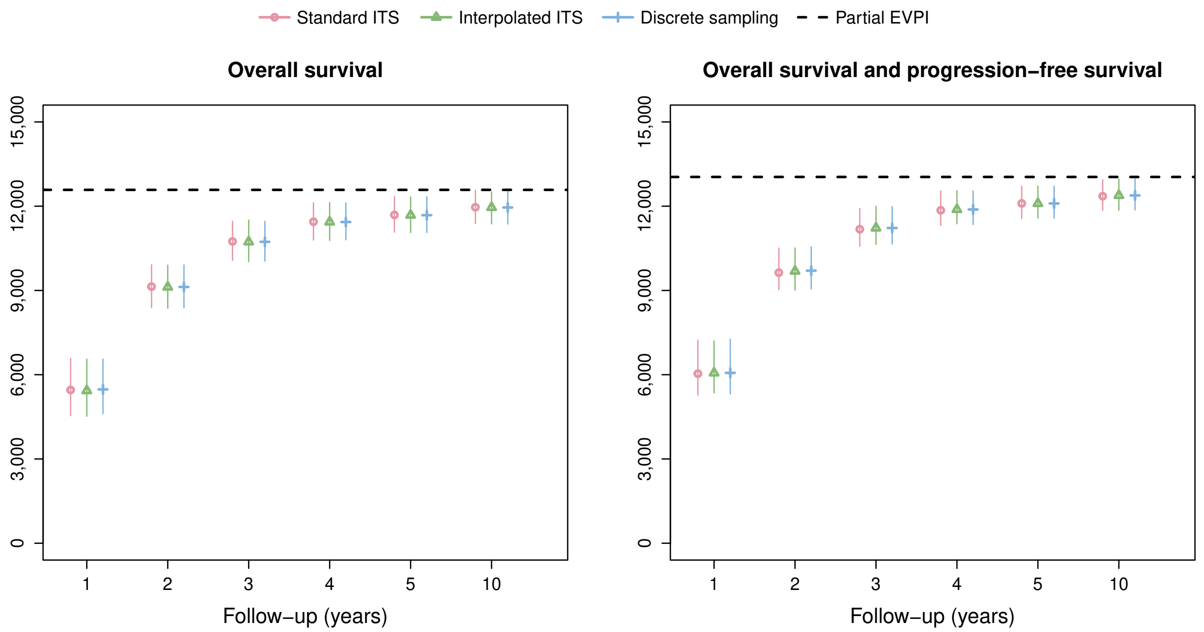


Figure 5: EVSI values for the synthetic case study with adjustment for treatment effect waning. Total computation times for the data-simulation procedures are 120,180 s (standard ITS), 33 s (interpolated ITS) and 12 s (discrete sampling).

Discussion

Strengths and limitations

We developed an interpolated ITS method and a discrete sampling method for simulating survival data from a probabilistic sample of survival probabilities over discrete time units. Our general-purpose methods greatly reduce the computational burden of the standard ITS method when closed-form expressions for the quantile function are unavailable, such as for custom distributions that incorporate assumptions about treatment effect waning as commonly encountered in practice[11], and for flexible survival models, including relative survival models, spline models, mixture cure models and response-based landmark models[9,10]. The implementation of our methods is identical across all possible survival models and can therefore easily be standardized from standard probabilistic decision analyses.

Generally, the precision of the EVSI estimator is influenced by the number of simulated datasets and the effective sample size of the simulated data. The discrete sampling method and, to a lesser degree, the interpolated ITS method, additionally introduce an approximation error that depends on the cycle length. It is generally recommended that discrete-time health economic models use a short cycle length to reduce the discrete-time approximation error, which could be as short as 1 week for slowly-progressing chronic diseases.[32] Our synthetic case study suggests that the approximation error introduced by our general-purpose methods is very small even when using a longer cycle length of 1 month in combination with short follow-up times and a low effective sample size of the simulated data given a rapidly-progressing disease.

We structured the synthetic case study around a PSM, a type of model that is frequently used to inform reimbursement decisions for new oncology drugs[33]. The key assumption behind a PSM is that survival endpoints, such as OS and PFS, are independent. This also implies that dependency between OS and PFS is not reflected in the EVSI data-simulation procedure when using a PSM. Joint modeling of OS and PFS could be implemented in a state transition model (STM), which uses transition probabilities to describe 'movements' between health states over time. STMs require individual patient data to estimate all relevant transition probabilities, unlike PSMs that can use digitised Kaplan-Meier data from published trials. OS and PFS data can be simulated jointly from a STM by first sampling a PFS time, and then deciding whether the sampled PFS time is a progression or death event using a binomial experiment with probability derived from the hazards of transitioning from PFS to PPS and OS.[34] If the sampled PFS time is a progression event, residual time until death can be simulated using the survival distribution for PPS to OS. Since transition probabilities are typically derived from survival curves fitted to time-to-event data, our data-simulation methods could also be useful in a STM framework.

If individual patient data are available with a similar study design and at least the same length of follow-up as the proposed study, study datasets could alternatively be simulated using a two-level resampling method based on bootstrapping[35]. In this approach, the observed dataset is first resampled K times with replacement, and then n values are sampled with replacement from the $k = 1, ..., K$ resampled datasets, thereby generating K new datasets with n observations each. In most situations, particularly for novel treatments, this type of data will however not be available.

The key notion behind EVSI is that the prior distribution of the model parameters is updated with simulated study data in order to estimate the joint posterior distribution given both prior information and the simulated study data. In EVSI analyses, the prior distribution is often informed by external evidence, such as digitised Kaplan-Meier data. This may however not match the way in which real world analyses of study data are conducted, since these may not synthesize the collected study data and the external evidence. Analysts should therefore ensure that the way in which the study data is analyzed once it has been collected is aligned with the assumptions underpinning the EVSI analysis.

Despite its routine application in cost-effectiveness analyses, there is currently a lack of guidance on how to model treatment effect waning.[12] In the synthetic case study, we modeled treatment effect waning by specifying probability distributions for the start and duration of the waning period, while preserving uncertainty about independent survival endpoints using a novel additive hazard approach. This had a large impact on the EVSI estimates, which highlights the importance of appropriately incorporating uncertainty about treatment effect waning in the EVSI calculations. There may be other possible approaches to model treatment effect waning, and these can easily be captured by our data-simulation methods as well.

Conclusion

The increasing prevalence of immature survival data in decision making, particularly for new cancer treatments[5], has been accompanied by the introduction of increasingly complex approaches for extrapolation[9,10], which complicates the EVSI data-simulation step. Our general-purpose data-simulation methods greatly reduce the computational burden of the EVSI data-simulation step when custom distributions that incorporate treatment effect waning or flexible survival models are used for which closed-form expressions for the quantile function are unavailable. Our methods are straightforward to implement and can easily be automated from standard probabilistic decision analyses, such as those used in technology assessments of new pharmaceuticals.[11,36–38] This means that our general-purpose methods can be used to simulate survival data - with a similar accuracy and computational cost - as using the correct closed-form quantile function for any survival model. Efficient EVSI calculations for survival data can help decision makers determine whether current evidence is sufficient or whether there is a need for collecting additional survival data before making an adoption decision.

Data Availability

The analysis code used in this study is available from Github at https://github.com/matverv/sim-surv-data-evsi.

Declaration of Conflicting Interests

The authors have no conflicts of interest that are directly relevant to the content of this article.

Appendices

Appendix A - Methods for Simulating Survival Times

In this appendix we will illustrate how to simulate survival times in R using the 3 methods discussed in the paper: standard inverse transform sampling (standard ITS), interpolated inverse transform sampling (interpolated ITS) and discrete sampling.

We will first install (if not installed yet) and load the flexsurv, tidyverse and knitr packages, which we will use in the example.

```
if ("flexsurv" %in% rownames(installed.packages()) == FALSE) {
  install.packages("flexsurv") # install flexsurv if not installed yet
}
if ("tidyverse" %in% rownames(installed.packages()) == FALSE) {
  install.packages("tidyverse") # install tidyverse if not installed yet
}
if ("knitr" %in% rownames(installed.packages()) == FALSE) {
  install.packages("knitr") # install flexsurv if not installed yet
}
library(flexsurv) # load flexsurv
library(tidyverse) # load tidyverse
library(knitr) # load knitr
```

Probabilistic Analysis

We start by running a probabilistic analysis in which we sample $k = 1, \ldots, K$ values $\boldsymbol{\theta}^{(k)}$ from the prior distribution of the model parameters, $p(\boldsymbol{\theta})$, which in this example is the rate parameter of an exponential model that describes overall survival for a single treatment arm.

```
set.seed(1) # set the seed for reproducibility
K <- 1000 # number of simulations
theta <- rgamma(K, 35, 950) # prior distribution for the exponential rate
```

We then construct K vectors of survival probabilities by evaluating the exponential survivor function given the sampled values $\boldsymbol{\theta}^{(k)}$ at discrete time cycles.

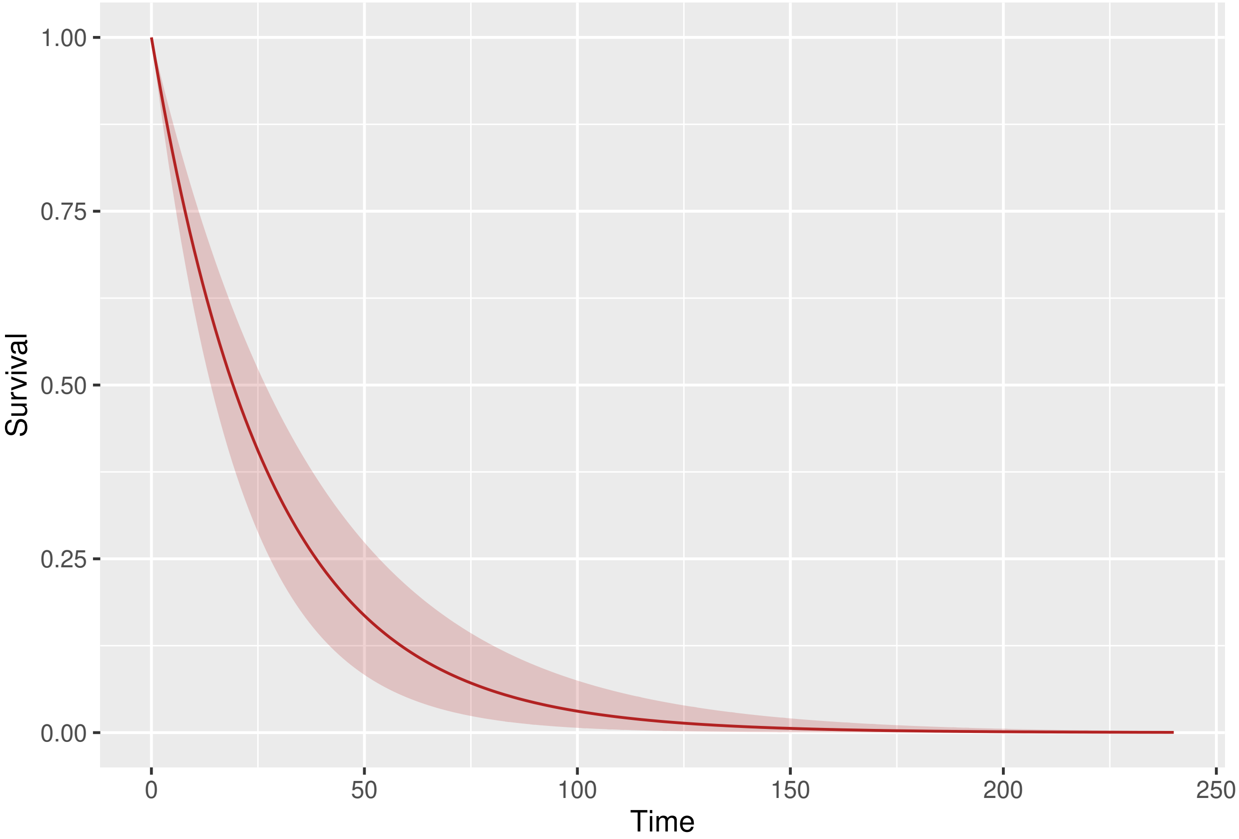
```
cycles <- 0:240 # discrete time cycles
surv <- sapply(theta, function (x) {
  pexp(cycles, x, lower.tail = F) # evaluate the survivor function
})
```

The following code produces a plot of the mean survival curve and 95% bounds.

```
# compute 95% percentiles
ci <- sapply(cycles+1, function (x) {quantile(surv[x,], c(0.025, 0.975))})

# create a dataframe of the sampled survival probabilities and confidence bounds
df_surv <- data.frame(
  "Time" = cycles,
  "Survival" = rowMeans(surv),
  "Upper" = ci[2, ],
  "Lower" = ci[1, ]
)

# plot mean survival and 95% confidence bounds
ggplot(df_surv, aes(Time, Survival)) +
  geom_ribbon(aes(ymin = Lower,
                  ymax = Upper),
                  fill = "firebrick",
                  alpha = 0.2) +
  geom_line(color = "firebrick")
```



Standard Inverse Transform sampling

Simulating Survival Data for a New Study

We can simulate survival times for a new study using the standard inverse transform sampling method by evaluating the quantile function at random uniform samples between 0 and 1. We can implement this using the `rexp` function in base `R`, which requires 2 arguments: the number of observations (e.g. number of patients to be included in the new study) and a value for the exponential rate parameter.

```
n <- 100 # number of patients/observations

# simulate n survival times given theta
times_std1 <- sapply(theta, function (x) rexp(n, x) )
```

The output is a matrix of simulated survival times, where rows index the n simulated survival times, and columns index the K simulations. The first 10 simulated survival times for the first 5 simulations are given by:

```
times_std1[1:10, 1:5]
##             [,1]      [,2]      [,3]      [,4]       [,5]
##  [1,] 22.5218924  3.943437  4.373529  8.961688  42.061293
##  [2,] 24.0482783  5.001096  8.227731 59.718875   9.319299
##  [3,]  0.7739475 32.194231  2.212844 16.903244  49.725231
##  [4,]  4.1427170 40.615923  5.390183 90.057154 142.905871
##  [5,] 39.9443093 52.405560 19.684164 62.166050   6.600515
##  [6,] 29.8874133 17.284868 18.631781  5.463462  26.606568
##  [7,] 22.5574594 20.743536  4.572290  5.475647 107.326405
##  [8,] 19.8110560 35.331819 38.729165 12.121705  51.730617
##  [9,]  6.4400761  4.422201 11.617468 42.017328  52.588304
## [10,] 10.5242127  1.511484 21.624664  9.374594   2.223138
```

Simulating Survival Data for an Ongoing Study

The `rexp` function simulates survival times that can take any value between 0 and infinity. If we want to simulate survival times for an ongoing study, we need to sample survival times that are greater than the observed follow-up time t_{obs}. We can do this by sampling from a conditional quantile function that is left-truncated at t_{obs}. We can implement this in `R` using the following code (we will record the computation time using the `Sys.time()` function):

```
# generate hypothetical observed follow-up times for the n patients
t_obs <- runif(n, 8, 12)

start_t <- Sys.time() # start the clock

# evaluate the survivor function at t_obs given theta
p_2 <- sapply(theta, function (x) pexp(t_obs, x, lower.tail = F))

# sample random numbers on the interval [0, p_2]
r_num <- sapply(1:K, function (x) runif(n, 0, p_2[,x]))

# evaluate the quantile function given theta at the random uniform values
times_std2 <- sapply(1:K, function (x) qexp(r_num[,x], theta[x], lower.tail = F))

ct_std2 <- difftime(Sys.time(), start_t,units = "secs") # stop the clock
```

Numerical Methods

When closed-form expressions for the quantile function, such as the `qexp` function, are unavailable, we could numerically evaluate the integrals and function inverses to produce a quantile function.

In the following code, we define a new exponential survivor function `f_surv` by first defining the hazard function, and then integrating the the hazard to produce the cumulative hazard. We can then compute the survival probabilities by exponentiating the negative cumulative hazards. The `f_surv(q, theta)` function numerically approximates the analytic exponential survivor function `pexp(q, rate=theta, lower.tail=FALSE)` in base `R`.

```
# exponential survivor function
f_surv <- function (q, theta) {
  f_haz <- hexp # exponential hazard function
  f_cum_haz <- function (q) { # cumulative hazard function
    integrate(f_haz, 0, q, rate = theta)$value
    }
  f_cum_haz <- Vectorize(f_cum_haz)
  cum_haz <- f_cum_haz(q)
  surv <- exp(-cum_haz) # survival probabilities
  return(surv)
}
```

We can invert the survivor function using the `qgeneric` function in `flexsurv` to produce the quantile function. We can then simulate survival times by evaluating the quantile function at a random uniform sample.

```
start_t <- Sys.time() # start the clock

# evaluate the survivor function at t_obs given theta
p_2 <- sapply(theta, function (x) f_surv(t_obs, x))

# sample random numbers on the interval [0, p_2]
r_num <- sapply(1:K, function (x) runif(n, 0, p_2[,x]))

# evaluate the quantile function given theta at the random uniform values
times_std3 <- sapply(1:K, function (x) {
  sapply(r_num[,x], function (y) qgeneric(f_surv, p = y, theta=theta[x]))
})

ct_std3 <- difftime(Sys.time(),start_t,units = "secs") # stop the clock
```

Interpolated Inverse Transform Sampling

We can also use inverse transform sampling to simulate survival times from the vectors of survival probabilities over discrete time units that we generated in the probabilistic analysis. Similar to the standard inverse transform sampling method, we first sample random uniform numbers. We then interpolate the survival probabilities using monotone cubic splines at the sampled numbers and record the interpolated cycle times

```
start_t <- Sys.time() # start the clock

# interpolate the model cycles at t_obs and record the survival probabilities
p_2 <- sapply(1:K, function (x) {
  spline(x=cycles, y=surv[,x], xout=t_obs, ties=max, method="hyman")$y
})

# sample random numbers on the interval [min(s), p_2]
r_num <- sapply(1:K, function (x) runif(n, min(surv[,x]) , p_2[,x]) )

# interpolate the survival probabilities at the random uniform numbers
# and record the corresponding cycle time
times_int <- sapply(1:K, function (x) {
  spline(x=surv[,x], y=cycles, xout=r_num[,x], ties=max, method="hyman")$y
})
```

```
ct_int <- difftime(Sys.time(),start_t,units = "secs") # stop the clock
```

Discrete Sampling

An alternative approach for simulating survival times from the vectors of survival probabilities over discrete time units is by sampling discrete ‘bins’ of cycle times with probability equal to the cumulative density of each bin. We can estimate the cumulative density of each bin by computing the successive differences between the survival probabilities at each discrete cycle time, which we then use as probabilities to sample the half-cycle times.

```
start_t <- Sys.time() # start the clock

# define the minimum and maximum half-cycle times
t_min <- round(mean(t_obs),0) + 0.5
t_max <- max(cycles-0.5)

# sample from the half-cycle times with probability equal to the
# successive differences between the survival probabilities at each cycle time
times_dis <- sapply(1:K, function (x) {
  sample(t_min:t_max, size = n,replace = TRUE,
         prob = abs(diff(surv[,x]))[ceiling(t_min):length(abs(diff(surv[,x])))])
})

ct_disc <- difftime(Sys.time(),start_t,units = "secs") # stop the clock
```

Results

Kernel Density Plot

The kernel density plot in Figure A1 shows that the distribution of the simulated survival times is very similar across the different methods.

```
df_times <- gather(as.data.frame(
  cbind(
    "Standard ITS (analytical)" = c(times_std2),
    "Standard ITS (numerical)" = c(times_std3),
    "Interpolated ITS" = c(times_int),
    "Discrete sampling" = c(times_dis)
  )
), factor_key = TRUE)

ggplot(df_times, aes(x=value, color = key)) +
  geom_density(alpha=0.4) +
  xlab("Simulated Survival Times") +
  ylab("Density") +
  labs(color = "Method")
```

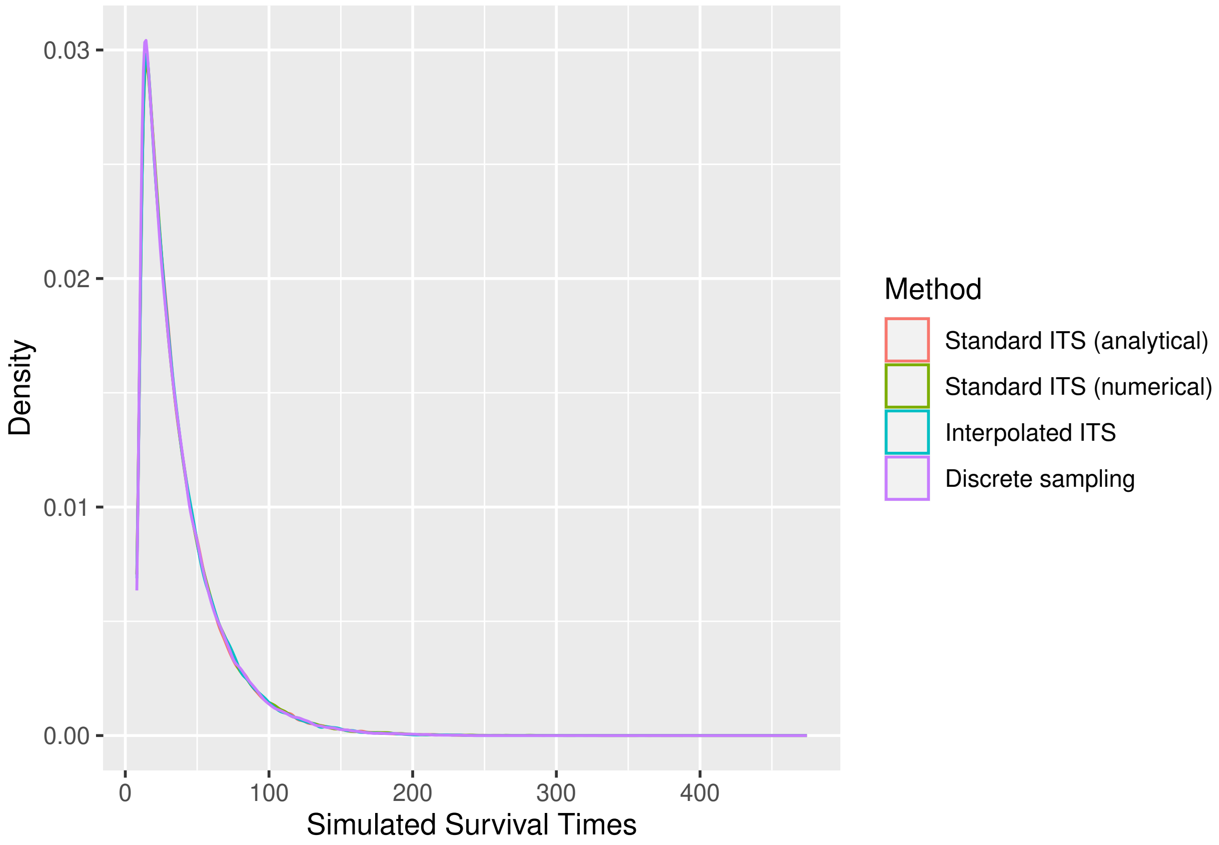


Figure A1: Density plot of the simulated survival times for the exponential example

Computation Time

The computation time for the different data-simulation procedures is given in Table A1. The standard inverse transform sampling scheme that uses numerical methods for evaluating the quantile function is several orders of magnitude slower than the other methods.

```
df_ct <- round(rbind(ct_std2, ct_std3, ct_int, ct_disc), 2)
rownames(df_ct) <- unique(df_times$key)
colnames(df_ct) <- "Computation time in seconds"
knitr::kable(df_ct,
             booktabs = TRUE,
             caption = "Computation times for the data-simulation procedures for the exponential example
             ) %>%
    add_footnote(" ",
                 threeparttable = T,
                 notation = "none")
```

Table A1: Computation times for the data-simulation procedures for the exponential example

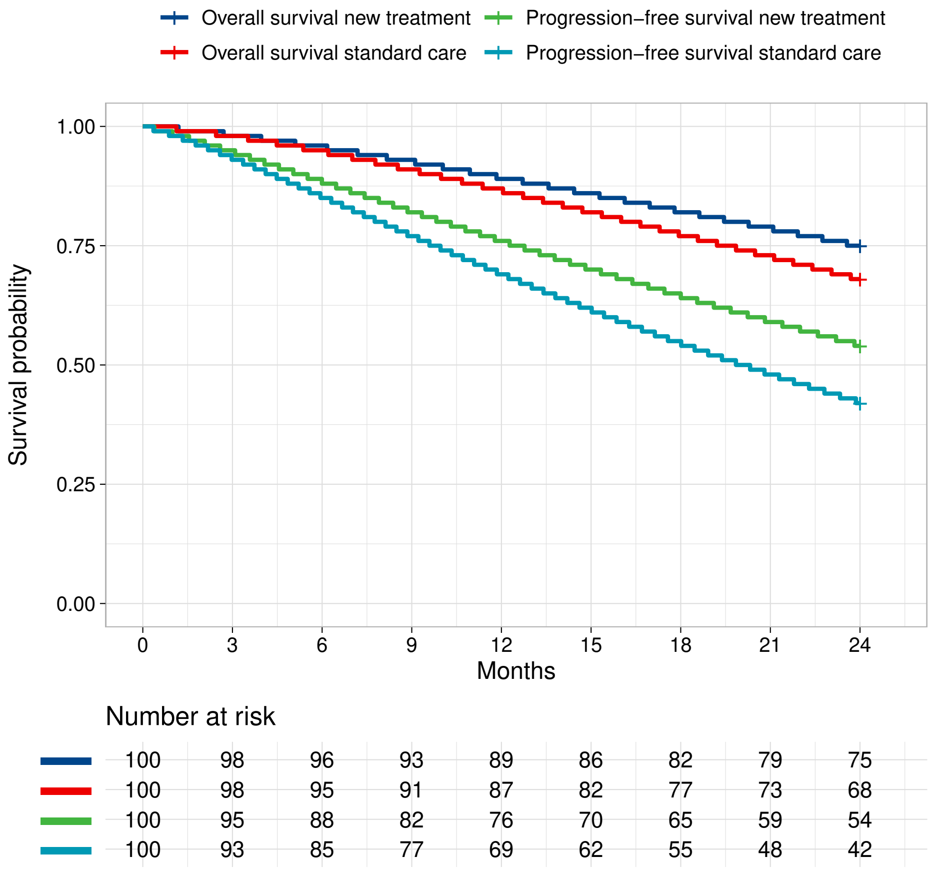
	Computation time in seconds
Standard ITS (analytical)	0.02
Standard ITS (numerical)	271.26
Interpolated ITS	0.35
Discrete sampling	0.03

Appendix B - Synthetic case study dataset

We generated a synthetic dataset using evenly spaced OS and PFS times for each d given 4 Weibull distributions, using the $0.005^{th}, 0.015^{th}, ..., 0.985^{th}, 0.995^{th}$ quantiles from each distribution. This avoids generating implausible datasets where PFS exceeds OS or OS for standard care exceeds OS for the new treatment, which could occur by randomly sampling survival times. We enrolled all patients at time zero and right censored the survival times at 24 months. We did not apply any other type of censoring. The parameters of the Weibull distributions that we used to generate the synthetic case study data set are shown in Table B1, and the Kaplan-Meier plot of the synthetic case study data set is given in Figure B1.

Table B1: Weibull distribution parameters for the synthetic case study dataset

	New treatment	Standard care
Overall survival		
Weibull log-shape, α	0.30	0.35
Weibull log-scale, β	4.10	3.85
Progression-free survival		
Weibull log-shape, α	0.15	0.20
Weibull log-scale, β	3.60	3.30



Number at risk

🟦	100	98	96	93	89	86	82	79	75
🟥	100	98	95	91	87	82	77	73	68
🟩	100	95	88	82	76	70	65	59	54
🟦	100	93	85	77	69	62	55	48	42

Figure B1: Kaplan-Meier plot for the synthetic case study dataset

Appendix C - Net benefit functions for the synthetic case study

The net benefit function for the new treatment is

$$
\begin{aligned}
\mathrm{NB}_1(\boldsymbol{\theta}) = \ & \frac{\lambda}{12} \sum_{c=0}^{C} e^{\left(-\left[\sqrt[12]{1+r}-1\right]\, t_c\right)} \left\{U_{pfs}\ S(t_c, \boldsymbol{\theta}_{pfs1}) + U_{pps}\ \left[S(t_c, \boldsymbol{\theta}_{os1}) - S(t_c, \boldsymbol{\theta}_{pfs1})\right]\right\} \\
& - \sum_{c=0}^{C} e^{\left(-\left[\sqrt[12]{1+r}-1\right]\, t_c\right)} \left\{C_{drug1}\ S(t_c, \boldsymbol{\theta}_{pfs1}) + C_{med1}\ S(t_c, \boldsymbol{\theta}_{os1})\right\},
\end{aligned}
\tag{17}
$$

and the net benefit for standard care is

$$
\begin{aligned}
\mathrm{NB}_2(\boldsymbol{\theta}) = \ & \frac{\lambda}{12} \sum_{c=0}^{C} e^{\left(-\left[\sqrt[12]{1+r}-1\right]\, t_c\right)} \left\{U_{pfs}\ S(t_c, \boldsymbol{\theta}_{pfs2}) + U_{pps}\ \left[S(t_c, \boldsymbol{\theta}_{os2}) - S(t_c, \boldsymbol{\theta}_{pfs2})\right]\right\} \\
& - \sum_{c=0}^{C} e^{\left(-\left[\sqrt[12]{1+r}-1\right]\, t_c\right)} \left\{C_{med2}\ S(t_c, \boldsymbol{\theta}_{os2})\right\},
\end{aligned}
\tag{18}
$$

where $S(t_c, \boldsymbol{\theta}_d)$ is the Weibull survivor function evaluated at cycle time t_c,

$$
S(t_c, \boldsymbol{\theta}_d) = e^{-\left(t_c/e^{\beta_d}\right)^{e^{\alpha_d}}}. \tag{19}
$$

References

1. Raiffa H, Schlaifer R. *Applied statistical decision theory.* Boston: Division of Research, Graduate School of Business Adminitration, Harvard University, 1961.
2. Vervaart M, Strong M, Claxton KP, et al. An Efficient Method for Computing Expected Value of Sample Information for Survival Data from an Ongoing Trial. *Medical Decision Making* 2022; 42: 612–625.
3. Heath A, Strong M, Glynn D, et al. Simulating Study Data to Support Expected Value of Sample Information Calculations: A Tutorial. *Medical Decision Making* 2022; 42: 143–155.
4. Latimer NR. Survival Analysis for Economic Evaluations Alongside Clinical Trials—Extrapolation with Patient-Level Data: Inconsistencies, Limitations, and a Practical Guide. *Medical Decision Making* 2013; 33: 743–754.
5. Tai T-A, Latimer NR, Benedict A, et al. Prevalence of Immature Survival Data for Anti-Cancer Drugs Presented to the National Institute for Health and Care Excellence and Impact on Decision Making. *Value in Health* 2020; S1098301520344648.
6. European Medicines Agency. Adaptive pathways. *European Medicines Agency*, https://www.ema.europa.eu/en/human-regulatory/research-development/adaptive-pathways (2018, accessed 17 July 2022).
7. European Medicines Agency. Conditional marketing authorisation. *European Medicines Agency*, https://www.ema.europa.eu/en/human-regulatory/marketing-authorisation/conditional-marketing-authorisation (2018, accessed 17 July 2022).
8. U.S. Food and Drug Administration (FDA). Fast Track, Breakthrough Therapy, Accelerated Approval, Priority Review. *FDA*, https://www.fda.gov/patients/learn-about-drug-and-device-approvals/fast-track-breakthrough-therapy-accelerated-approval-priority-review (2019, accessed 17 July 2022).
9. Rutherford MJ, Lambert PC, Sweeting MJ, et al. NICE DSU TECHNICAL SUPPORT DOCUMENT 21: Flexible Methods for Survival Analysis. 97.
10. Palmer S, Borget I, Friede T, et al. A Guide to Selecting Flexible Survival Models to Inform Economic Evaluations of Cancer Immunotherapies. *Value in Health*; 0.
11. National Institute for Health and Care Excellence. *Guide to the methods of technology appraisal 2013*, https://www.nice.org.uk/process/pmg9 (2013).
12. Witlox WJA, van Asselt ADI, Wolff R, et al. Durvalumab for the Treatment of Locally Advanced, Unresectable, Stage III Non-Small Cell Lung Cancer: An Evidence Review Group Perspective of a NICE Single Technology Appraisal. *Pharmacoeconomics* 2020; 38: 317–324.
13. Lambe T, Duarte R, Mahon J, et al. Cladribine Tablets for the First-Line Treatment of Relapsing-Remitting Multiple Sclerosis: An Evidence Review Group Perspective of a NICE Single Technology Appraisal. *PharmacoEconomics* 2019; 37: 345–357.
14. Strong M, Oakley JE, Brennan A, et al. Estimating the Expected Value of Sample Information Using the Probabilistic Sensitivity Analysis Sample: A Fast, Nonparametric Regression-Based Method. *Medical Decision Making* 2015; 35: 570–583.
15. Briggs A, Claxton K, Sculpher M. *Decision Modelling for Health Economic Evaluation.* 1 edition. Oxford: Oxford University Press, 2006.
16. Ades AE, Lu G, Claxton K. Expected value of sample information calculations in medical decision modeling. *Medical Decision Making* 2004; 24: 207–227.
17. Collett D. *Modelling Survival Data in Medical Research.* 3 edition. Chapman and Hall/CRC, 2015.
18. Latimer NR. Survival Analysis for Economic Evaluations Alongside Clinical Trials - Extrapolation with Patient-Level Data: Inconsistencies, Limitations, and a Practical Guide. *Medical Decision Making* 2013; 33: 743–754.
19. R Core Team. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/ (2021).

20. Jackson C, Stevens J, Ren S, et al. Extrapolating Survival from Randomized Trials Using External Data: A Review of Methods. *Medical Decision Making* 2017; 37: 377–390.

21. Kunst N, Wilson ECF, Glynn D, et al. Computing the Expected Value of Sample Information Efficiently: Practical Guidance and Recommendations for Four Model-Based Methods. *Value in Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research* 2020; 23: 734–742.

22. Latimer NR, Adler AI. Extrapolation beyond the end of trials to estimate long term survival and cost effectiveness. *BMJ Medicine*; 1.

23. Jackson CH, Thompson SG, Sharples LD. Accounting for uncertainty in health economic decision models by using model averaging. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 2009; 172: 383–404.

24. Jackson CH, Sharples LD, Thompson SG. Structural and parameter uncertainty in Bayesian cost-effectiveness models. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 2010; 59: 233–253.

25. Fragoso TM, Neto FL. Bayesian model averaging: A systematic review and conceptual classification. *International Statistical Review* 2018; 86: 1–28.

26. Akaike H. Information Theory and an Extension of the Maximum Likelihood Principle. In B. N. Petrov, & F. Csaki (Eds.), Proceedings of the 2nd International Symposium on Information Theory (pp. 267-281). *Budapest: Akademiai Kiado* 1973; p 267–81.

27. Woods BS, Sideris E, Palmer S, et al. Partitioned Survival and State Transition Models for Healthcare Decision Making in Oncology: Where Are We Now? *Value in Health* 2020; 23: 1613–1621.

28. Bojke L, Soares MO, Claxton K, et al. Reference Case Methods for Expert Elicitation in Health Care Decision Making. *Medical Decision Making* 2022; 42: 182–193.

29. Heath A, Manolopoulou I, Baio G. Efficient Monte Carlo Estimation of the Expected Value of Sample Information Using Moment Matching. *Medical Decision Making* 2018; 38: 163–173.

30. Wood S. Mgcv: Mixed GAM Computation Vehicle with Automatic Smoothness Estimation, https://CRAN.R-project.org/package=mgcv (2020).

31. Strong M, Oakley JE, Brennan A. Estimating Multiparameter Partial Expected Value of Perfect Information from a Probabilistic Sensitivity Analysis Sample: A Nonparametric Regression Approach. *Medical Decision Making* 2014; 34: 311–326.

32. O'Mahony JF, Newall AT, van Rosmalen J. Dealing with Time in Health Economic Evaluation: Methodological Issues and Recommendations for Practice. *PharmacoEconomics* 2015; 33: 1255–1268.

33. Woods BS, Sideris E, Palmer S, et al. Partitioned Survival and State Transition Models for Healthcare Decision Making in Oncology: Where Are We Now? *Value in Health* 2020; 23: 1613–1621.

34. Meller M, Beyersmann J, Rufibach K. Joint modeling of progression-free and overall survival and computation of correlation measures. *Statistics in Medicine* 2019; 38: 4270–4289.

35. Sadatsafavi M, Marra C, Bryan S. Two-Level Resampling as a Novel Method for the Calculation of the Expected Value of Sample Information in Economic Trials. *Health Economics* 2013; 22: 877–882.

36. Canada's Drug and Health Technology Agency (CADTH). Guidelines for the Economic Evaluation of Health Technologies: Canada (4th Edition), https://www.cadth.ca/sites/default/files/pdf/guidelines_for_the_economic_evaluation_of_health_technologies_canada_4th_ed.pdf (2017).

37. Commonwealth of Australia. Guidelines for preparing a submission to the Pharmaceutical Benefits Advisory Committee, Version 5.0, https://pbac.pbs.gov.au/content/information/files/pbac-guidelines-version-5.pdf (2016).

38. European Network for Health Technology Assessment. Practical considerations when critically assessing economic evaluations, https://www.eunethta.eu/wp-content/uploads/2020/03/EUnetHTA-JA3WP6B2-5-Guidance-Critical-Assessment-EE_v1-0.pdf (2020).