Calculating the Expected Value of Sample Information using Efficient Nested Monte Carlo: A Tutorial - Supplementary Material

Appendix A. Code for the Toy Example

This section of code allows you to replicate the "Toy" Examples given in the main paper.

```
################################################################################
#Calculating the EVSI for the Toy Example                                      #
################################################################################
#Run set.seed(1) to get the same results as those given in the paper
set.seed(1)

#PSA samples for the 3 parameters
e1<-rbeta(10,3,4)
e2<-rbeta(10,4,3)
c<-rnorm(10,3,20)

#Calculating the INB
INB<-100*(e1-e2)-c
#Saving the mean and variance of the INB
mu.theta<-mean(INB)
sigma.theta<-var(INB)

#Value of eliminating ALL model uncertainty

#The pmax function takes the max between 0 and each INB separately
EVPI<-mean(pmax(0,INB))-max(0,mean(INB))

#Finding the value of reducing ALL the uncertainty for each parameter

#The gam function fits the non-parameteric regression - see Strong et al. (2014)
library(mgcv)
#e1
gam.e1<-gam(INB~e1)
(mean(pmax(0,gam.e1$fitted))-max(0,mean(gam.e1$fitted)))/EVPI
#e2
gam.e2<-gam(INB~e2)
(mean(pmax(0,gam.e2$fitted))-max(0,mean(gam.e2$fitted)))/EVPI
#c
gam.c<-gam(INB~c)
(mean(pmax(0,gam.c$fitted))-max(0,mean(gam.c$fitted)))/EVPI

#e1 is the most valuable - so we will calculate the EVSI of a sample targeting e1
```

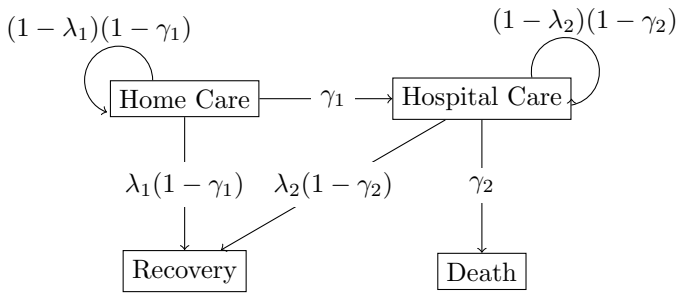
```
#Save the fitted value for e1
fitted.phi<-(gam.e1$fitted)
#Find the variance of the fitted values
sigma.phi<-var(gam.e1$fitted)

#Sampling the "future samples" X for Q=3
#The function quantile gives the q/(Q+1) quantiles for e1
Q<-3
quantile(e1,probs=1:Q/(Q+1),type=4)
set.seed(1)
X1<-rbinom(1,20,0.285)
X2<-rbinom(1,20,0.47)
X3<-rbinom(1,20,0.52)

#Finding the variances by sampling from the posterior of e1|X
#NOTE: This is usually done by posterior updating but we have a conjugate model
#and a linear INB so we can use conjugacy. In standard modelling this would need to
#be done using MCMC.

set.seed(1)
e1.X1<-rbeta(100,3+X1,4+20-X1)
INB.X1<-100*(e1.X1-e2)-c
S1<-var(INB.X1)
e1.X2<-rbeta(100,3+X2,4+20-X2)
INB.X2<-100*(e1.X2-e2)-c
S2<-var(INB.X2)
e1.X3<-rbeta(100,3+X3,4+20-X3)
INB.X3<-100*(e1.X3-e2)-c
S3<-var(INB.X3)

#The variances for the different samples
S1;S2;S3
#Take the mean variance to calculate the EVSI
sigma.X<-mean(c(S1,S2,S3))

#Calculating the scaled INB to calculate the EVSI
INB.scale<-(fitted.phi-mu.theta)/sqrt(sigma.phi)*sqrt(sigma.theta-sigma.X)+mu.theta

#Calculating the EVSI
#The function pmax() takes the maximum between 0 and each INB.scale
mean(pmax(0,INB.scale))-max(0,mu.theta)
#EVSI calculation as give in the paper
1/10*(29+2+10+4)-max(0,mu.theta)

#############################################################################
#NOTE: The EVSI using this code will be slightly different to the value in the    #
#paper as the values are rounded in the paper so there is some loss of accuracy   #
#############################################################################
```

Appendix B. A Health Economic Model to Evaluate A New Chemotherapy Treatment

The model evaluates the economic efficiency of two chemotherapy treatments, a standard of care ($t = 0$) and a novel treatment ($t = 1$) that aims to reduce the number of side effects. It is assumed that the two treatments are equally effective in terms of treating cancer and therefore the model focuses solely on the treatment of side effects. In this regard, we define π_0 as the probability of experiencing side effects on the standard of care and ρ percentage change in the number of side effects for the treatment. It is assumed that we have access to data from a previous study where the standard of care was given to 111 patients of which 27 experienced side effects. The prior for ρ is defined from a literature review, with the distribution defined in Table B.1.

If a patient does not experience side effects then they incur no additional cost and are associated with a QALY q which has a PSA distribution with mean 0.98 and variance 0.001. Evidently, every treated patient will incur a drug cost of either £110 for the standard of care or £420 for the novel treatment. These costs are known with certainty.

To model the costs and QALYs associated with experiencing side effects, we model the progression of the side effects using a 4 state Markov Model. Once a patient experiences side effects it is assumed that the disease progression is the same irrespective of the Chemotherapy treatment they received. Therefore, for all patients experiencing side effects we use the following model to calculate the costs and QALYs,



All transitions that are not marked on this diagram are assumed to have probability 0. This implies that the progression of side effects is based on 4 model inputs, γ_1 and γ_2 which give the probability of requiring hospital care and the probability of dying respectively and λ_1 and λ_2 which give the probability of recovery given that you remain in home care or hospital care respectively.

It is assumed that we have access to data from our previous study, stating that of the 27 patients who experienced side effects, 17 required hospital care and 1 died. This gives indirect information about γ_1 and γ_2. Specifically, we assume a cycle length of 1 day for the Markov Model with a 15 day time horizon. We assume that γ_1 and γ_2 remain constant across the time horizon and therefore define the model inputs $\Gamma_1 = 15\gamma_1$ and $\Gamma_2 = 15\gamma_2$ which are then directly informed by the available data. On the other hand, PSA distributions for λ_1 and λ_2 are defined using informative beta distributions with means 0.45 and 0.35 respectively and variance 0.02. It is assumed that patients in the state Recovery have the same QALY as patients who did not experience side effects and no additional cost.

Finally, to determine the cost and QALY of experiencing side effects, we assigned a cost and QALY score to each state. Firstly, we assumed that patients in Recovery had the same profile as patients who did not experience side effects. Secondly, if a patient died, they incurred a cost of terminal care c_{death} once with a QALY score of 0. Both Home Care and Hospital Care were given a cost, c_{HC} and c_H respectively, and a QALY score, q_{HC} and q_H. In this case, hospital treatment was more costly than home care and has a lower QALY score. For all these model inputs, we use lognormal PSA distributions for costs and beta distributions for QALYs.

Appendix B.1. The Sampling Distributions to Inform $\boldsymbol{\phi}$

In the main paper, we calculated the EVSI for a trial designed to inform 6 key parameters $\boldsymbol{\phi} = (\pi_0, \rho, \gamma_1, \gamma_2, \lambda_1, \lambda_2)$. To inform these six parameters, we intend to collect six outcomes from the future

Model Input	Distribution	1st Prior Parameter	2nd Prior Parameter	Previous Data
π_0	Beta	1	1	Number of side effects
ρ	Normal	Mean: 0.65	Var: 0.01	No
q	Beta	18.23	0.372	No
Γ_1	Beta	1	1	Number of hospitalizations
Γ_2	Beta	1	1	Number of deaths
γ_1	$\frac{\Gamma_1}{15}$	-	-	-
γ_2	$\frac{\Gamma_2}{15}$	-	-	-
λ_1	Beta	5.12	6.26	No
λ_2	Beta	3.63	6.74	No
c_{death}	LogNormal	8.33	0.13	No
c_{HC}	LogNormal	7.74	0.039	No
c_H	LogNormal	8.77	0.15	No
q_{HC}	Beta	5.75	5.75	No
q_H	Beta	0.87	3.47	No

Table B.1: The full prior specification for the model inputs for the Chemotherapy example presented in the main paper. This table gives the distributional assumption and the parameters given for that parameter, unless specified the prior parameters are given in the order they would be in the JAGS modelling language for Bayesian updating. We also indicate whether the prior distributions are combined with addition data before calculating the EVSI.

trial. Firstly, we collect the number of side effects on each arm of the trial, assumed to follow a binomial distribution conditional on π_0 and ρ;

$$X_{SE0} \sim Binomial(150, \pi_0)$$

$$X_{SE1} \sim Binomial(150, \rho\pi_0).$$

Conditional on the number of patients who experience side effects, we then investigate the number of patients who receive hospital treatment and the number of patients who die;

$$X_{Hosp} \sim Binomial(X_{SE0} + X_{SE1}, \gamma_1)$$

$$X_{Death} \sim Binomial(X_{Hosp}, \gamma_2).$$

Finally, the recovery times for patients who receive home care and patients who receive hospital care are assumed to follow exponential distributions conditional on a function of the transition probabilities. Specifically, we take

$$\eta_1 = -\log(\lambda_1)$$

and

$$\eta_2 = -\log(\lambda_2),$$

as parameters to model

$$T_{HC} \sim Exponential(\lambda_1),$$

the time to recover for each patient who only receive home care, therefore X_{HC} is a vector of length $X_{SE0} + X_{SE1} - X_{Hosp}$ and

$$T_H \sim Exponential(\lambda_2),$$

the time to recover for patients who receive hospital treatment, making X_H a vector of length $X_{Hosp} - X_{Death}$.

Clearly, this model specification induces non-linearities in the model structure and non-conjugacy for the parameters. Therefore, this example requires the use of nested Monte Carlo, especially as the model structure also induces dependence between the parameters informed by the study.

Appendix C. A brief outline of the theory underpinning the Heath *et al.* method

Recall that the method is based on approximating

$$\mathrm{E}_{\boldsymbol{\theta}|\boldsymbol{X}}\left[\mathrm{INB}\right]$$

using a small number of samples for $\boldsymbol{X}$. To understand the method, we need to think more about this expectation. In normal Bayesian analysis, this expectation is a fixed number because the data have been observed, so you have a true posterior and therefore the posterior expectation for the INB is easy to calculate.

However, in this analysis, the data hasn't been collected and therefore have a distribution which gives the likelihood of all the potential future samples of $\boldsymbol{X}$. For each of these potential samples the posterior mean will be slightly different and therefore we have a distribution over the potential future posterior means for the INB. This implies that *before* seeing the data, the posterior mean could be equal to a large number of different possibilities and it's only once we observe the data that we know the value of the actual posterior mean.

The Moment Matching method uses the PSA simulations of the INB to estimate a sample from the distribution of the future posterior means; $\mathrm{INB}^s_{\boldsymbol{\phi}}$. Specifically, the PSA simulations for the INB, conditional only on the parameters that are being directly updated by the sample information, are rescaled so they are similar to the distribution of the future posterior means. Heath *et al.* [**?**] show that the distribution of the future posterior means is similar to $\mathrm{INB}^s_{\boldsymbol{\phi}}$ in most practical settings, especially when the sample size of the future data set is sufficiently large > 30.

In fact, it is possible to demonstrate that the main difference between the distribution of the future posterior mean and $\mathrm{INB}^s_{\boldsymbol{\phi}}$ is the *variance* of the two distributions. Therefore, as long as the variance of the future posterior mean is known, we can rescale $\mathrm{INB}^s_{\boldsymbol{\phi}}$ to have this variance and we achieve a good estimation for the distribution of interest.

However, as $\mathrm{E}_{\boldsymbol{\theta}|\boldsymbol{X}}\left[\mathrm{INB}\right]$ is a conditional expectation, there is a formula available to calculate the variance:

$$\begin{aligned}\mathrm{Var}_{\boldsymbol{X}}\left[\mathrm{E}_{\boldsymbol{\theta}|\boldsymbol{X}}\left[\mathrm{INB}\right]\right] &= \mathrm{Var}_{\boldsymbol{\theta}}\left[\mathrm{INB}_{\boldsymbol{\theta}}\right] - \\ &\mathrm{E}_{\boldsymbol{X}}\left[\mathrm{Var}_{\boldsymbol{\theta}|\boldsymbol{X}}\left[\mathrm{INB}_{\boldsymbol{\theta}}\right]\right] \\ &= \sigma_{\boldsymbol{\theta}} - \sigma_{\boldsymbol{X}}.\end{aligned}$$

This formula implies that, theoretically, this variance is equal to the variance of the prior INB (the variance of the PSA simulations) minus the mean posterior variance (estimated from Q posteriors).

We can use Q posterior distributions to estimate $\sigma_{\boldsymbol{X}}$ because $\mathrm{E}_{\boldsymbol{X}}\left[\mathrm{Var}_{\boldsymbol{\theta}|\boldsymbol{X}}\left[\mathrm{INB}_{\boldsymbol{\theta}}\right]\right]$ can be rewritten as

$$\mathrm{E}_{\boldsymbol{\phi}}\left[\mathrm{E}_{\boldsymbol{X}|\boldsymbol{\phi}}\left[\mathrm{Var}_{\boldsymbol{\theta}|\boldsymbol{X}}\left[\mathrm{INB}_{\boldsymbol{\phi}}\right]\right]\right].$$

Choosing Q quantiles for $\boldsymbol{\phi}$ allows us to calculate the outer expectation as efficiently as possible spanning all the possible values of $\boldsymbol{\phi}$, a method known as quadrature. Finally, sampling $\boldsymbol{X}$ conditional on the value of $\boldsymbol{\phi}$ calculates the inner expectation by simulation, which is why Q should be greater than 30.

Appendix D. Extension to multi-decision

In a multi-decision setting, we consider a model with T alternative or novel treatment options being compared with the standard of care. In this setting therefore, we have T different INBs — the difference between the novel treatment option and the standard of care. Therefore, performing PSA implies that we generate S simulations of the INB for each novel treatment option. To distinguish the INB for each treatment option we will use the notation INB^t.

To calculate the EVSI in the multi-decision setting, we need:

- The fitted values $\mathrm{INB}^t_{\boldsymbol{\phi}}$, for $t = 1, \ldots, T$, i.e. T different sets of S fitted values obtained while calculating the EVPPI or using non-parametric regression.

- The mean of $\text{INB}^t_{\boldsymbol{\theta}} = \mu^t_{\boldsymbol{\theta}}$ for every $t = 1, \ldots, T$, these should be stored in a *vector* which is called $\mu_{\boldsymbol{\theta}}$ in the formula below.
- The variance of INB is also needed. As the INB now has multiple treatment options, the "variance" of the INB is a $T \times T$ sample variance-covariance matrix, where the covariance between the different INB values is contained in the off-diagonal entries in the matrix. To denote a matrix, we use $\Sigma_{\boldsymbol{\theta}}$.
- The variance-covariance matrix is also needed for the fitted values, which we denote $\Sigma_{\boldsymbol{\phi}}$. Note that in `R` the variance-covariance matrix is easily calculated using the `var()` function on matrix of INB values.
- Finally, the mean variance-covariance matrix, denoted $\Sigma_{\boldsymbol{X}}$ must be estimated from the hypothetical posterior simulations in the same way as for the duel-decision setting. This implies that you need to choose evenly spaced $\boldsymbol{\phi}$ values to generate $\boldsymbol{X}$ and then update the posterior for the model inputs using $\boldsymbol{X}$. These model inputs are then used to find the new INB for each treatment. The variance-covariance matrix then needs to be calculated for these updated INB values.

Once all these elements is available, the EVSI is estimated by rescaling the INB for all the different treatment options. This needs to be done in one step using the variance-covariance matrices.

$$\text{INB}^* = \left(\text{INB}^{\boldsymbol{\phi}} - \mu^{\boldsymbol{\theta}}\right) \left(\Sigma^{\boldsymbol{\phi}}\right)^{-\frac{1}{2}} \left(\Sigma^{\boldsymbol{\theta}} - \Sigma^{\boldsymbol{X}}\right)^{\frac{1}{2}} + \mu^{\boldsymbol{\theta}},$$

where, the power $\frac{1}{2}$ is the matrix square root of the variance-covariance matrix and $-\frac{1}{2}$ is its inverse. Standard statistical software makes this matrix algebra relatively simple, for example in `R` there is a `sqrtm` function in the `Matrix` package available to find these matrices.

[1] H. Tuffaha, L. Gordon, and P. Scuffham. Value of information analysis informing adoption and research decisions in a portfolio of health care interventions. *MDM Policy & Practice*, 1(1):2381468316642238, 2016.

[2] M. Strong, J. Oakley, A. Brennan, and P. Breeze. Estimating the expected value of sample information using the probabilistic sensitivity analysis sample a fast nonparametric regression-based method. *Medical Decision Making*, page 0272989X15575286, 2015.