Simulation Sample Sizes for Monte Carlo Partial EVPI Calculations

Jeremy E. Oakley,

Department of Probability and Statistics, University of Sheffield, UK

Alan Brennan, Paul Tappenden and Jim Chilcott

School of Health and Related Research, University of Sheffield, UK.

November 10, 2009

Running title: Partial EVPI sample sizes.

Address for Correspondence: Dr Jeremy Oakley, Department of Probability and Statistics, University of Sheffield, The Hicks Building, Hounsfield Road, Sheffield, S3 7RH, UK. E-mail: j.oakley@sheffield.ac.uk

Summary

Partial expected value of perfect information (EVPI) quantifies the economic value of removing uncertainty concerning parameters of interest in a decision model. EVPIs can be computed via Monte Carlo methods. An outer loop samples values of the parameters of interest, and for each of these, an inner loop samples the remaining parameters from their conditional distribution. This nested Monte Carlo approach can result in biased estimates if very small numbers of inner samples are used and can require a large number of model runs for accurate partial EVPI estimates. We present a simple algorithm to estimate the EVPI bias and EVPI confidence interval for a specified number of inner and outer samples. The algorithm uses a relatively small number of model runs (we suggest approximately 600), is quick to compute, and can help determine how many outer and inner iterations for a desired level of accuracy. This algorithm is tested using three case studies: two illustrative cost-effectiveness models of different complexity, with easily computed partial EVPIs, and a much more complex health economic model of multiple sclerosis. The first two case studies demonstrate that our algorithm produces robust estimates of the EVPI bias and EVPI confidence interval for different numbers of inner and outer samples. In the complex case study, such large Monte Carlo sample sizes are required that estimation is not feasible due to computational time and alternative computational methods are required. This algorithm to establish numbers of model runs required for partial EVPI estimation will be of use in any decision model when parameter uncertainty is important.

KEY WORDS: Economic Model, Expected Value of Perfect Information, Monte Carlo estimation, Bayesian Decision Theory.

1 Introduction

It is common practice to use decision models to estimate the expected net benefit of alternative strategy options open to a decision maker (Raiffa, 1968), (Brennan and Akehurst, 2000). Invariably in health economic cost-effectiveness models, the input parameters' true values are not known with certainty. A probabilistic sensitivity analysis will then be required to investigate the consequences of this input parameter uncertainty (Briggs and Gray, 1999). Simple Monte Carlo propagation of input uncertainty through the model can provide an estimate of the distribution of net benefit, thus giving its expected value, and the probability that the incremental net benefit will be positive (Van Hout et al., 1994).

More detailed analysis can compute the partial expected value of perfect information (partial EVPI), that is, the value to the decision maker of learning the true value of the uncertain parameter input before deciding whether to adopt the new treatment (Raiffa, 1968), (Claxton and Posnett, 1996). Partial EVPIs are recommended because they quantify the importance of different parameters using decision-theoretic arguments (Claxton, 1999), (Meltzer, 2001), and thus can quantify the societal value of further data collection to help design and prioritise research projects (Claxton and Thompson, 2001), (Chilcott et al., 2003a). Unfortunately, evaluating partial EVPIs is computationally demanding. Generally, a two level Monte Carlo

procedure is needed, requiring many runs of the economic model. A developing literature has described this more and more clearly (Felli and Hazen, 1998), (Brennan et al., 2002), (Felli and Hazen, 2003), (Ades et al., 2004), (Yokota and Thompson, 2004), (Koerkamp et al., 2006), (Brennan et al., 2007), (Brennan and Kharroubi, 2007b). The procedure begins with an outer loop sampling values of the parameters of interest, and for each of these, an inner loop sampling the remaining parameters from their conditional distribution (Brennan et al., 2002) (Ades et al., 2004).

Increasingly large inner and outer sample sizes will produce increasingly accurate estimates of partial EVPI. However, the nested Monte Carlo approach can result in biased estimates if the inner loop sample size is small (Brennan and Kharroubi, 2007a), regardless of the outer loop sample size. Most studies recommend 'large enough' inner and outer samples e.g. 1,000 or 10,000 in order to produce 'accurate' partial EVPI estimates, but none discuss in detail the choice of sample sizes in relation to bias or confidence interval widths.

In this paper we present an algorithm for determining the inner and outer loop sample sizes needed to estimate a partial EVPI to a required level of accuracy. We first review the theory of estimating partial EVPIs via Monte Carlo. We propose using normal approximations for conditional expected net benefits to estimate the bias and variance of a Monte Carlo estimate for a specified number of inner and outer loops. We set out an algorithm, using a moderate number of model runs, to produce an estimate of the bias and confidence interval widths for partial EVPI estimates, and hence to help determine the inner and outer sample sizes needed. The

results of applying and testing the algorithm's performance in three case studies are given, followed by a discussion of how this approach can be applied more generally to support Monte Carlo estimation of partial EVPI.

If the decision model is complex, requiring substantial computation time for each model run, then it may not be feasible to do the number of model runs required for our algorithm, or the algorithm may suggest that an infeasible number of model runs are required for sufficiently accurate partial EVPI estimates. In this case, one can use instead a Gaussian process meta-model, which approximates the economic model, enabling more efficient Monte Carlo sampling, but with the disadvantages that it is far more complex to program initially and is not always feasible for models with large numbers of uncertain input parameters (e.g., in excess of 100) (Oakley, 2009), (Oakley and O'Hagan, 2004), (Stevenson et al., 2004)

2 Methods

2.1 Evaluating Partial EVPI via Monte Carlo Sampling

Suppose we have T treatment options, and our economic model computes the net benefit $NB(t, \mathbf{x})$ for treatment $t = 1, \ldots, T$, when provided with input parameters $\mathbf{x}$. We denote the true, uncertain values of the input parameters by $\mathbf{X} = \{X_1, \ldots, X_d\}$, so that the true, uncertain net benefit of treatment t is given by $NB(t, \mathbf{X})$. When considering the partial EVPI of a particular parameter X_i, we use the notation $\mathbf{X}_{-i} = \{X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_d\}$ to denote all the inputs in $\mathbf{X}$ except X_i. For

any input X, we use subscripts to denote a particular input parameter (or group of parameters) in the economic model, and a superscript to denote a randomly sampled value of that input. Note that all of the equations presented, and indeed the proposed algorithm, are the same if we are considering a group of parameters $\mathbf{X}_i$ rather than a single scalar parameter.

We use the notation $NB(t, \mathbf{X}) = NB(t, X_i, \mathbf{X}_{-i})$ and for expectations $E_{\mathbf{X}}$, E_{X_i} and $E_{\mathbf{X}_{-i}|X_i}$ denote expectations over the full joint distribution of $\mathbf{X}$, the marginal distribution of X_i, and the conditional distribution of $\mathbf{X}_{-i}|X_i$ respectively.

The partial EVPI for the ith parameter X_i is given by

$$EVPI(X_i) = E_{X_i}\left[\max_t E_{\mathbf{X}_{-i}|X_i}\left\{NB(t, X_i, \mathbf{X}_{-i})\right\}\right] - \max_t E_{\mathbf{X}}\{NB(t, \mathbf{X})\}. \quad (1)$$

The second term in the RHS of (1) can be estimated by Monte Carlo. We sample $\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(N)}$ from the distribution of $\mathbf{X}$, evaluate $NB(t, \mathbf{X}^{(n)})$ for $t = 1, \ldots, T$ and $n = 1, \ldots, N$, and then estimate $NB^* = \max_t E_{\mathbf{X}}\{NB(t, \mathbf{X})\}$ by

$$\widehat{NB}^* = \max_t \frac{1}{N} \sum_{n=1}^{N} NB(t, \mathbf{X}^{(n)}).$$

We do not consider the choice of N in this paper. In practice, it is usually feasible to have N sufficiently large such that $\widehat{NB}^*$ is an accurate estimate of NB^*, and can be used in the partial EVPI estimate of any parameter or group of parameters.

In this paper we concentrate on estimating the first term in the RHS of (1). We define the maximum conditional expected net benefit given X_i as

$$m(X_i) = \max_t E_{\mathbf{X}_{-i}|X_i}\left\{NB(t, X_i, \mathbf{X}_{-i})\right\},$$

and hence, the first term on the RHS of equation (1) can be written as $E_{X_i}\{m(X_i)\}$.

Typically, we cannot evaluate $m(X_i)$ analytically, and so we estimate it using Monte Carlo. We randomly sample J values of $\mathbf{X}_{-i}$ from the conditional distribution of $\mathbf{X}_{-i}|X_i$ to obtain $\mathbf{X}_{-i}^{(1)}, \ldots, \mathbf{X}_{-i}^{(J)}$. We run the economic model for each of the J sampled inputs to obtain $NB(t, X_i, \mathbf{X}_{-i}^{(j)})$ for $t = 1, \ldots, T$ and $j = 1, \ldots, J$, and estimate $m(X_i)$ by

$$\hat{m}(X_i) = \max_t \frac{1}{J} \sum_{j=1}^{J} NB(t, X_i, \mathbf{X}_{-i}^{(j)}). \tag{2}$$

We now approximate $E_{X_i}\{m(X_i)\}$ by $E_{X_i}\{\hat{m}(X_i)\}$, and estimate $E_{X_i}\{\hat{m}(X_i)\}$ by randomly sampling K values $X_i^{(1)}, \ldots, X_i^{(K)}$ from the distribution of X_i and computing the estimator

$$\hat{E}_{X_i}\{\hat{m}(X_i)\} = \frac{1}{K} \sum_{k=1}^{K} \hat{m}\{X_i^{(k)}\}. \tag{3}$$

We refer to the process of obtaining the maximum conditional net benefit estimator $\hat{m}\{X_i\}$ for a single given X_i using J samples from the distribution of $\mathbf{X}_{-i}|X_i$ as the *inner level* Monte Carlo procedure. The process of calculating (3) (given the values $\hat{m}(X_{i,1}), \ldots, \hat{m}(X_{i,K})$) using the K samples $X_i^{(1)}, \ldots, X_i^{(K)}$ is the *outer level* Monte Carlo procedure. The Monte Carlo estimate for the partial EVPI for parameter X_i is therefore

$$\begin{aligned} \widehat{EVPI}(X_i) &= \hat{E}_{X_i}\{\hat{m}(X_i)\} - \widehat{NB}^* \\ &= \frac{1}{K} \sum_{k=1}^{K} \left[\max_t \left\{ \frac{1}{J} \sum_{j=1}^{J} NB\left(t, X_i^{(k)}, \mathbf{X}_{-i}^{(j,k)}\right) \right\} \right] \\ &\quad - \max_t \frac{1}{N} \sum_{n=1}^{N} NB(t, \mathbf{X}^{(n)}), \end{aligned} \tag{4}$$

where $\mathbf{X}_{-i}^{(j,k)}$ is the jth sample from the distribution of $\mathbf{X}_{-i}|X_i = X_i^{(k)}$. The inner and outer level sample sizes, J and K respectively, mean that the total number of

runs of the economic model required for the partial EVPI estimate is $J \times K$ (given $\widehat{NB}^*$). The objective in this paper is to determine what values of J and K should be used in order to obtain a sufficiently accurate estimate of the partial EVPI.

2.2 Uncertainty and Bias in Monte Carlo Estimates of Partial EVPI

Any Monte Carlo estimate of an expectation is subject to uncertainty due to random sampling, and the larger the number of samples, the more this uncertainty is reduced. We assume that N is sufficiently large such that uncertainty in NB^* is small relative to uncertainty in $E_{X_i}\{m(X_i)\}$. If K is sufficiently large then a normal approximation will apply, and a 95% confidence interval for $EVPI(X_i)$ is given by

$$\left(\widehat{EVPI}(X_i) - 1.96\sqrt{\frac{Var\{\hat{m}(X_i)\}}{K}},\ \widehat{EVPI}(X_i) + 1.96\sqrt{\frac{Var\{\hat{m}(X_i)\}}{K}}\right). \qquad (5)$$

This interval appears to suggest that to obtain a sufficiently accurate estimate of the partial EVPI, we just need K to be large, as the width of the confidence interval will decrease as K increases. Unfortunately, this is not the case because, as we now show, $\hat{m}(X_i)$ is an upwards biased estimator of $m(X_i)$, and so $\widehat{EVPI}(X_i)$ is a biased estimator of $EVPI(X_i)$. The bias is independent of K, but it depends on J, and can be reduced to an acceptably small level if we choose J sufficiently large.

We define the conditional expected net benefit given a particular X_i for each treatment t as

$$\mu_t(X_i) = E_{\mathbf{X}_{-i}|X_i}\{NB(t, X_i, \mathbf{X}_{-i})\}$$

and its estimator based on J random samples as

$$\hat{\mu}_t(X_i) = \frac{1}{J}\sum_{j=1}^{J} NB(t, X_i, \mathbf{X}_{-i}^{(j)}).$$

In our estimate of EVPI in equation (4), we estimate the maximum conditional expected net benefit over the T treatments given a particular X_i,

$$m(X_i) = \max\{\mu_1(X_i), \ldots, \mu_T(X_i)\},$$

by the maximum of the Monte Carlo estimates for each treatment

$$\hat{m}(X_i) = \max\{\hat{\mu}_1(X_i), \ldots, \hat{\mu}_T(X_i)\}.$$

The problem is that, although $\hat{\mu}_t(X_i)$ is an unbiased estimator of $\mu_t(X_i)$, i.e. $E_{\mathbf{X}_{-i}|X_i}\{\hat{\mu}_t(X_i)\} = \mu_t(X_i)$, when the maximisation is applied to the estimators then $\hat{m}(X_i)$ is not an unbiased estimator of $m(X_i)$. It is upwards biased i.e. it tends to over-estimate $m(X_i)$. This is because for any set of random variables $Z_1, \ldots, Z_n$, it is straightforward to show that $E\{\max(Z_1, \ldots, Z_n)\} \geq \max\{E(Z_1), \ldots, E(Z_n)\}$, which follows from Jensen's inequality. Applying this to the estimator $\hat{m}(X_i)$, we have

$$\begin{aligned}
E_{\mathbf{X}_{-i}|X_i}\{\hat{m}(X_i)\} &= E_{\mathbf{X}_{-i}|X_i}\left[\max\{\hat{\mu}_1(X_i), \ldots, \hat{\mu}_T(X_i)\}\right] \\
&\geq \max\left[E_{\mathbf{X}_{-i}|X_i}\{\hat{\mu}_1(X_i)\}, \ldots, E_{\mathbf{X}_{-i}|X_i}\{\hat{\mu}_T(X_i)\}\right] \\
&= \max\{\mu_1(X_i), \ldots, \mu_T(X_i)\} \\
&= m(X_i).
\end{aligned}$$

That is, $E_{\mathbf{X}_{-i}|X_i}\{\hat{m}(X_i)\} \geq m(X_i)$. Thus, we expect $\hat{m}(X_i)$ to overestimate $m(X_i)$ for any value of X_i and consequently we expect $\widehat{EVPI}(X_i)$ to overestimate the true partial EVPI.

2.3 A multivariate normal approximation for conditional expected net benefit estimators

To quantify the bias in the estimator $\hat{m}(X_i)$ we can investigate the sampling distribution (the probability distribution, under repeated sampling of the population) of the vector of Monte Carlo estimators for conditional expected net benefit $\hat{\boldsymbol{\mu}}(X_i) = \{\hat{\mu}_1(X_i), \ldots, \hat{\mu}_T(X_i)\}'$. If J is sufficiently large then, a multivariate normal approximation with T dimensions will apply. Thus we have

$$\hat{\boldsymbol{\mu}}(X_i) \sim N_T \left\{ \boldsymbol{\mu}(X_i), \frac{1}{J} V(X_i) \right\}, \tag{6}$$

where $\boldsymbol{\mu}(X_i) = \{\mu_1(X_i), \ldots, \mu_T(X_i)\}'$ and element p, q of the matrix $V(X_i)$ is given by

$$V_{p,q}(X_i) = Cov_{\mathbf{X}_{-i}|X_i} \left\{ \hat{\mu}_p(X_i), \hat{\mu}_q(X_i) \right\}.$$

The bias, $E_{\mathbf{X}_{-i}|X_i}\{\hat{m}(X_i)\} - m(X_i)$, will depend on $\boldsymbol{\mu}(X_i)$, $V(X_i)$ and J, and will decrease as J increases.

We illustrate this with a simple example. Suppose we have two treatments $T = 2$, with

$$\hat{\boldsymbol{\mu}}(X_i) = \begin{pmatrix} \hat{\mu}_1(X_i) \\ \\ \hat{\mu}_2(X_i) \end{pmatrix} \sim N_2 \left\{ \begin{pmatrix} 9000 \\ \\ 10000 \end{pmatrix}, \frac{1}{J} \begin{pmatrix} 1500^2 & 475^2 \\ \\ 475^2 & 1500^2 \end{pmatrix} \right\}.$$

In figure 1 we plot the density functions of $\hat{\mu}_1(X_i)$ and $\hat{\mu}_2(X_i)$ for $J = 1$ and $J = 100$. When $J = 1$, there is considerable overlap between the two densities, with the result that $E[\max\{\hat{\mu}_1(X_i), \hat{\mu}_2(X_i)\}]$ (marked as the vertical line) is noticeably greater than $\max[E\{\hat{\mu}_1(X_i)\}, E\{\hat{\mu}_2(X_i)\}] = 10000$. When $J = 100$, the overlap is reduced and $E[\max\{\hat{\mu}_1(X_i), \hat{\mu}_2(X_i)\}]$ is much closer to 10000, i.e. the bias is negligible.

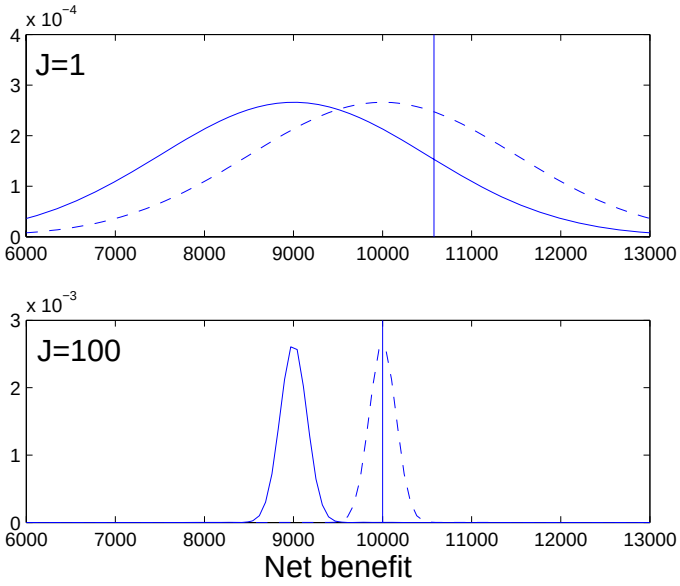

Figure 1: The sampling distributions of $\hat{\mu}_1(X_i)$ (solid line) and $\hat{\mu}_2(X_i)$ (dashed line) for $J = 1$ and $J = 100$. The vertical line shows the expectation of the maximum of $\hat{\mu}_1(X_i)$ and $\hat{\mu}_2(X_i)$ in each case. The bias is expectation of the maximum minus the maximum of the two expectations, 10000.

3 Estimating the bias and confidence interval width in $\widehat{EVPI}(X_i)$ for a specified J and K

We now present an algorithm to estimate the bias and confidence interval width for a specified J and K. An example code listing for implementing the algorithm in R (R Development Core Team, 2006) can be downloaded from `www.jeremy-oakley.staff.shef.ac.uk/evpissize.R`. (The precise implementation of our algorithm will depend on the nature of the economic model and the input distributions)

There are three stages to the algorithm. In the first stage we run the economic model a moderate number of times (630 is proposed) to obtain estimates of $\boldsymbol{\mu}(X_i)$ and $V(X_i)$, for different values of X_i. In the second stage, we use these estimates to sample from the distribution of the conditional expected net benefit estimators $\hat{\boldsymbol{\mu}}$ and estimate the bias in $\widehat{EVPI}(X_i)$ for different specified J's. In the third stage, we estimate the variance in the maximum conditional expected net benefits $Var\{\hat{m}(X_i)\}$, enabling us to estimate the width of a (95%) confidence interval for $\widehat{EVPI}(X_i)$ using equation (5) for different specified K's.

Stage 1: Estimating $\boldsymbol{\mu}(X_i)$ and $V(X_i)$

1. Draw a small random sample of points (say 21) $X_i^{(1)}, \ldots, X_i^{(21)}$ from the distribution of X_i.

2. For each $X_i^{(k)}$ generate a random sample (say 30) of the remaining inputs from their conditional distribution $\mathbf{X}_{-i}|X_i^{(k)}$. Denote these by $\mathbf{X}_{-i}^{(1,k)}, \ldots, \mathbf{X}_{-i}^{(30,k)}$.

3. Run the economic model to obtain the net benefits $NB(t, X_i^{(k)}, \mathbf{X}_{-i}^{(j,k)})$ for $t = 1, \ldots, T$; $j = 1, \ldots, 30$; $k = 1, \ldots, 21$.

4. Estimate $\mu_t(X_i^{(k)})$ by

$$\tilde{\mu}_t(X_i^{(k)}) = \frac{1}{30} \sum_{j=1}^{30} NB(t, X_i^{(k)}, \mathbf{X}_{-i}^{(j,k)}),$$

for $t = 1, \ldots, T$; $k = 1, \ldots, 21$ to obtain

$$\tilde{\boldsymbol{\mu}}(X_i^{(k)}) = \left\{\tilde{\mu}_1(X_i^{(k)}), \ldots, \tilde{\mu}_T(X_i^{(k)})\right\},$$

for $k = 1, \ldots, 21$.

5. Estimate $V(X_i^{(k)})$ for $k = 1, \ldots, 21$ by $\tilde{V}(X_i^{(k)})$ where element p, q of $\tilde{V}(X_i^{(k)})$ is given by

$$\frac{1}{29}\sum_{j=1}^{30}\left\{NB(p, X_i^{(k)}, \mathbf{X}_{-i}^{(j,k)}) - \tilde{\mu}_p(X_i^{(k)})\right\}\left\{NB(q, X_i^{(k)}, \mathbf{X}_{-i}^{(j,k)}) - \tilde{\mu}_q(X_i^{(k)})\right\}$$

We now have estimates of $\boldsymbol{\mu}(X_i)$ and $V(X_i)$ at a range of different values of X_i, and can now do a simulation to identify a suitable value of J.

Stage 2: Estimating bias and determining J

1. Choose a candidate value of J.

2. For $k = 1, \ldots, 21$:

   (a) Approximate the distribution of $\hat{\boldsymbol{\mu}}(X_i^{(k)})$ by

$$\hat{\boldsymbol{\mu}}(X_i^{(k)}) \sim N_T\left\{\tilde{\boldsymbol{\mu}}(X_i^{(k)}), \frac{1}{J}\tilde{V}(X_i^{(k)})\right\}$$

   (b) Generate $\hat{\boldsymbol{\mu}}^{(1)}(X_i^{(k)}), \ldots, \hat{\boldsymbol{\mu}}^{(N)}(X_i^{(k)})$ for large N (say 10,000) from the distribution of $\hat{\boldsymbol{\mu}}(X_i^{(k)})$, with

$$\hat{\boldsymbol{\mu}}^{(n)}(X_i^{(k)}) = \left\{\hat{\mu}_1^{(n)}(X_i^{(k)}), \ldots, \hat{\mu}_T^{(n)}(X_i^{(k)})\right\}',$$

   for $n = 1, \ldots, N$.

   (c) Estimate the bias in $\hat{m}(X_i^{(k)})$ for each of the 21 sampled $X_i^{(k)}$ by

$$\hat{b}(X_i^{(k)}) = \frac{1}{N}\sum_{n=1}^{N}\max\left\{\hat{\mu}_1^{(n)}(X_i^{(k)}), \ldots, \hat{\mu}_T^{(n)}(X_i^{(k)})\right\} - \max\left\{\tilde{\mu}_1(X_i^{(k)}), \ldots, \tilde{\mu}_T(X_i^{(k)})\right\} \tag{7}$$

3. Estimate the expected bias in $\hat{m}(X_i)$ by

$$\hat{b}(X_i) = \frac{1}{21}\sum_{k=1}^{21}\hat{b}(X_i^{(k)}). \tag{8}$$

This process is repeated for different J to determine the relationship between J and the scale of bias. When put into the context of the overall EVPI and the partial estimate $\widehat{EVPI}(X_i)$, this can help determine a J which produces an acceptably small bias.

Finally, we consider estimating the width of the confidence interval for the estimator in (3) that will result from choosing an outer sample size K.

Stage 3: Estimating the 95% CI for $\widehat{EVPI}(X_i)$ to determine K

There are two sources of variation in the estimator $\widehat{EVPI}(X_i)$ given in equation (4). The first is due to variation in the sample $X_i^{(1)}, \ldots, X_i^{(K)}$. The second is due to $\hat{m}(X_i^{(k)})$ being evaluated using a second Monte Carlo simulation: for a given $X_i^{(k)}$ the variance of $\hat{m}(X_i^{(k)})$ will be a function of the inner sample size J. We write

$$\hat{m}(X_i) = m(X_i) + \varepsilon(X_i, J),$$

where $\varepsilon(X_i, J)$ is the random error in the estimate $\hat{m}(X_i)$ of $m(X_i)$ based on a Monte Carlo sample of size J. We now propose an approximate estimate of $Var\{\hat{m}(X_i)\}$ based on two simplifications. Firstly, we suppose that $m(X_i)$ and $\varepsilon(X_i, J)$ are independent. Secondly, we ignore variation in $\varepsilon(X_i, J)$ due to variation in X_i. Specifically, we consider an average variance of $\varepsilon(X_i, J)$, averaging across X_i. We write

$$Var\{\hat{m}(X_i)\} = Var\{m(X_i)\} + Var\{\varepsilon(X_i, J)\}.$$

In our case studies, we find that for moderate J the variance of $\hat{m}(X_i)$ is dominated by the first term $Var\{m(X_i)\}$, suggesting that the two simplifications are reasonable.

We estimate $Var\{m(X_i)\}$ as follows.

1. Compute the average of the maximum conditional net benefits by

$$\bar{m}(X_i) = \frac{1}{K}\sum_{k=1}^{K} \hat{m}(X_i^{(k)}),$$

2. Estimate $Var\{m(X_i)\}$ by

$$\widehat{Var}\{m(X_i)\} = \frac{1}{K-1}\sum_{k=1}^{K}\left\{\hat{m}(X_i^{(k)}) - \bar{m}(X_i)\right\}^2. \tag{9}$$

We estimate $Var\{\varepsilon(X_i, J)\}$ by considering the variance of the Monte Carlo error for each $X_{(i)}^k$, and then taking the average for $k = 1, \ldots, 21$:

1. For each $X_i^{(k)}$, obtain the sample variance of the bias terms in (7):

$$\widehat{Var}\left\{b(X_i^{(k)})\right\} = \frac{1}{N-1}\sum_{n=1}^{N}\left\{b^{(n)}(X_i^{(k)}) - \hat{b}(X_i^{(k)})\right\}^2,$$

with $\hat{b}(X_i^{(k)})$ given in equation (7), and

$$b^{(n)}(X_i^{(k)}) = \max\left\{\hat{\mu}_1^{(n)}(X_i^{(k)}), \ldots, \hat{\mu}_T^{(n)}(X_i^{(k)})\right\} - \max\left\{\tilde{\mu}_1(X_i^{(k)}), \ldots, \tilde{\mu}_T(X_i^{(k)})\right\}.$$

2. Estimate $Var\{\varepsilon(X_i, J)\}$ by

$$\widehat{Var}\{\varepsilon(X_i, J)\} = \frac{1}{K}\sum_{k=1}^{K}\widehat{Var}\left\{b(X_i^{(k)})\right\}$$

Finally, we estimate the width of a 95% confidence interval for any value of J and K as

$$2 \times 1.96 \times \sqrt{\frac{\widehat{Var}\{m(X_i)\} + \widehat{Var}\{\varepsilon(X_i, J)\}}{K}}$$

3.1 Discussion

The most important purpose of the algorithm is to estimate the bias and determine J. This is because the size of the bias cannot be observed when we come to actually estimate a partial EVPI using Monte Carlo sampling. Although it is useful at stage 3 to estimate the width of the confidence interval for $\widehat{EVPI}(X_i)$, it is possible to observe directly the variability in EVPI estimates as more outer samples are used in the two-level procedure.

The algorithm uses the normal approximation based on the central limit theorem in both the confidence interval (5) and the distribution of $\boldsymbol{\mu}$ in (6). We would expect this approximation to be sufficient for practical purposes in most cases. The model user will want to find J and K such the bias and confidence interval width will be acceptably small, and so the it will usually be sufficient to estimate the bias and confidence interval width to within the correct order of magnitude only; very precise estimates should not be necessary.

In stage 1 of the algorithm, we have to sample from the conditional distribution of $\mathbf{X}_{-i}|X_i$. While this is straightforward for independent inputs, or suitably tractable families of multivariate distributions, such sampling may be computationally intensive for more complex joint distributions. For example, we may need to use Markov chain Monte Carlo to sample from the conditional distribution. Though this would make the algorithm harder to implement, it would of course make calculating a partial EVPI more computationally demanding in any case. In particular, we emphasise that our algorithm only requires the same sampling procedures used

to compute the partial EVPI itself (plus routine sampling from multivariate normal distributions).

Alternatives to simple random sampling can be used in stage 1, step 1. If X_i is a scalar, we could choose evenly spaced points and use an appropriate numerical integration procedure such as Simpson's rule as an alternative to the Monte Carlo estimates in (8) in (9). For example, we would instead calculate

$$\hat{b}(X_i) = \sum_{k=1}^{21} w_k f_{X_i}(X_i^{(k)}) \hat{b}(X_i^{(k)}),$$

where $f_{X_i}(.)$ is the marginal density of X_i, and w_k is a weight determined by the corresponding numerical integration rule. Another alternative is to use a stratified random sample, for example, Latin Hypercube sampling if X_i is a group of inputs.

4 Case studies

Three case studies have been used to explore the feasibility, accuracy and usefulness of the proposed algorithm in predicting the bias and confidence interval widths in EVPI estimates using a specified J and K.

4.1 Case study 1: A simple decision tree model

The first case study concerns a simple hypothetical cost-effectiveness model, which compares two strategies: treatment with drug T0 versus treatment with drug T1. Table 1 shows the nineteen uncertain model parameters, with prior mean values shown for T0 (column a), T1 (column b) and hence the incremental analysis (column

c). Costs include cost of drug and cost of hospitalisations, that is, the product of the percentage of patients admitted to hospital, days in hospital, and cost per day. Thus, mean cost of strategy $T0$ = \$1000 + 10% × 5.20 × \$400 = \$1208. Health benefits are measured as QALYs gained and come from two sources: responders receive a utility improvement for a specified duration, and some patients have side effects with a utility decrement for a specified duration i.e. QALY for strategy $T0 = 70\% \times \text{responders} \times 0.3 \times 3\text{years} + 25\%\text{side effects} \times -0.1 \times 0.5\text{years} = 0.6175$. The illustrative willingness to pay i.e. threshold cost per QALY is set at λ_w = \$10000. (This is a purely illustrative model; the value of λ_w was arbitrarily chosen and is lower than the threshold of many western countries). Thus, the net benefit of T0 is \$10000 × 0.6175 − \$1208 = \$4967. Effectively this is a simple decision tree model with a net benefit function of sum-product form i.e.

$$\begin{aligned} NB(T0) &= \lambda_w(X_5X_6X_7 + X_8X_9X_{10}) - (X_1 + X_2X_3X_4), \\ NB(T1) &= \lambda_w(X_{14}X_{15}X_{16} + X_{17}X_{18}X_{19}) - (X_{11} + X_{12}X_{13}X_4). \end{aligned}$$

The uncertain parameters are characterised with independent normal distributions. Standard deviations for the model parameters are shown in columns (d) and (e). Each parameter can be informed by collection of further data on individual patients. Given current knowledge, the basic model results show \$5405 expected net benefit for T1 compared with \$4967 for T0 (difference = \$437.80), which means that our baseline decision given current information should be to adopt strategy T1. Probabilistic sensitivity analysis (Briggs and Gray, 1999) shows that T1 provides greater net benefits than T0 on only 54.5% of 1000 Monte Carlo samples. This suggests

that knowing more about some of the uncertain parameters could affect our decision. The results of the partial EVPI analyses we have undertaken are described on an indexed scale, where the overall expected value of perfect information, \$1319 per patient in our model, is indexed to 100.

Our analysis of the accuracy of the algorithm in predicting the expected bias and confidence interval of EVPI estimates focusses on parameter X_5, the % responding to treatment $T0$. For inner sample sizes of J=10, 100 and 500 and outer samples also of K= 10, 100 and 500 we undertake comparisons between the predictions of our algorithm and the results of re-running the Monte-Carlo estimation of EVPI 1000 different times. By generating 1000 real estimates of $EVPI(X_5)$, using for example $J = 10$ inner and $K = 10$ outer samples, we can compute the mean bias exhibited and the 95% confidence interval exhibited in these EVPI estimates. We then compare the mean bias exhibited with that predicted by the algorithm and similarly the confidence interval width exhibited with that predicted by the algorithm.

4.2 Case study 2: A three state Markov cost-effectiveness model over 20 periods

The second case-study is an extension of case study 1, whereby the parameter regarding the duration of response to each individual drug is replaced with a more complex Markov transition process, which involves probabilities of moving between states responding, non-responding and death over a 20 cycle time horizon. Un-

	a	b	c	d	e
Parameters	**Parameter Mean Values given Existing Evidence**			**Uncertainty in Means Standard Deviations**	
	T0	T1	(T1-T0)	T0	T1
Cost of Drug (X_1, X_{11})	\$ 1000	\$1500	\$500	\$1	\$1
% Admissions (X_2, X_{12})	10%	8%	-2%	2%	2%
Days in Hospital (X_3, X_{13})	5.20	6.10	0.90	1.00	1.00
Cost per day (X_4)	\$400	\$400	...	\$200	\$200
% Responding (X_5, X_{14})	70%	80%	10%	10%	10%
Utility Change if respond (X_6, X_{15})	0.30	0.30	...	0.10	0.05
Duration of respose (years) (X_7, X_{16})	3.0	3.0	...	0.5	1.0
%Side effects (X_8, X_{17})	25 %	20 %	-5%	10%	5%
Change in utility if side effect (X_9, X_{18})	-0.10	-0.10	-0.00	0.02	0.02
Duration of side effect (years) (X_{10}, X_{19})	0.50	0.50	...	0.20	0.20
Central Estimate Results					
Total Cost	\$1208	\$1695	487		
Total QALY	0.6175	0.7100	0.0925		
Cost per QALY			5267		
Net Benefit given threshold (λ)= \$10000)	\$4967	\$5405	\$438		

Table 1: Summary of Model and Parameters for Case Study 1

certainty regarding the two separate transition matrices for the two treatments is modelled using Dirichlet distributions. Table 2 shows the distributions for all model parameters and the mathematical descriptions for the two net benefit functions.

This case study is more complex than case study 1. It has substantial non-linearity due to the matrix multiplication of the transition matrices over the twenty cycles. Uncertainty in model parameters in no longer described simply by independent normal distributions because the Dirichlet distributions are used, and hence, correlation between the transition matrix parameters is now inherent. Based on a probabilistic sensitivity analysis using 1 million simulations, treatment T1 is marginally better than T0 (expected net benefits of \$6164.898 versus \$6019.102) but there remains substantial uncertainty with the overall EVPI estimated at \$2045.431.

In this case study, the analysis of the accuracy of our algorithm for predicting the confidence interval and bias for estimated partial EVPI calculations focusses on the partial EVPI for the subset of parameters associated with the transition matrices i.e. parameter set X_{20} to X_{31}.

4.3 Case study 3: A complex cost-effectiveness model

The third case study uses a published health economic model developed on behalf of the National Institute for Clinical Excellence (NICE) to evaluate the cost-effectiveness of disease-modifying therapies (interferon-beta 1a, interferon-beta 1b and glatiramer acetate) versus conventional treatment in the management of multiple sclerosis in the UK. Full details are given elsewhere (Chilcott et al., 2003b), but

Parameters	**Treatment T0**			**Treatment T1**		
		Prior Mean	s		Prior Mean	s
Normally Distributed Parameters						
Cost of Drug ($)	X_1	1000	1	X_{11}	1500	1
Probability of admissions	X_2	0.10	0.02	X_{12}	0.08	0.02
Days in Hospital	X_3	5.20	1.00	X_{13}	6.10	1.00
Cost per day ($)	X_4	400	200	X_4	400	200
Probability of initial response	X_5	0.70	0.10	X_{14}	0.80	0.10
Utility Change if respond	X_6	0.30	0.10	X_{15}	0.30	0.05
Probability of side effects	X_8	0.25	0.10	X_{17}	0.20	0.05
Change in utility if side effect	X_9	-0.10	0.02	X_{18}	-0.10	0.02
Duration of side effect (years)	X_{10}	0.50	0.20	X_{19}	0.50	0.20
Markov Transition Matrices and Probabilities						
p(responding → responding)	X_{20}	0.60	Dirichlet	X_{26}	0.60	Dirichlet
p(responding → not responding)	X_{21}	0.30	(7,4,2)	X_{27}	0.30	(7,4,2)
p(responding → die)	X_{22}	0.10		X_{28}	0.10	
p(not resp.→ responding)	X_{23}	0	Dirichlet	X_{26}	0	Dirichlet
p(not resp.→ not responding)	X_{24}	0.90	(1,10,2)	X_{27}	0.90	(1,10,2)
p(not resp.→ die)	X_{25}	0.10		X_{28}	0.10	
p(die → die)	X_{22}	1.00		X_{28}	1.00	
Mathematical Relationships and Functions						
	M0$=\begin{pmatrix} X_{21} & X_{21} & X_{22} \\ X_{23} & X_{24} & X_{25} \\ 0 & 0 & 1 \end{pmatrix}$			M1$=\begin{pmatrix} X_{26} & X_{27} & X_{28} \\ X_{29} & X_{30} & X_{315} \\ 0 & 0 & 1 \end{pmatrix}$		
	$X_{22} = 1 - X_{21} - X_{20}$			$X_{28} = 1 - X_{27} - X_{26}$		
	$X_{25} = 1 - X_{24} - X_{23}$			$X_{31} = 1 - X_{30} - X_{29}$		
	$S0 = (X_5, 1 - X_5, 0)^T$			$S1 = (X_{14}, 1 - X_{14}, 0)^T$		
	$U0 = (X_6, 0, 0)^T$			$U1 = (X_{15}, 0, 0)^T$		
$NBT0 = \lambda \times \left\{\sum_{p=1}^{20} \left(S0^T \times (M0)^p \times U0\right) + X_8 + X_9 + X_{10}\right\} - (X_1 + X_2 \times X_3 \times X_4)$						
$NBT1 = \lambda \times \left\{\sum_{p=1}^{20} \left(S1^T \times (M1)^p \times U1\right) + X_{17} + X_{18} + X_{19}\right\} - (X_{11} + X_{12} \times X_{13} \times X_{14})$						

Table 2: **Summary of Model and Parameters for Case Study 2**

an overview of the model is as follows.

The model uses the state transition methodology to describe the clinical course of patients with relapsing/remitting MS (RRMS) and secondary progressive MS (SPMS) through the Expanded Disability Status Scale (EDSS). The frequency of MS relapse is superimposed upon each EDSS state. Instantaneous hazard rates are used to model the progression of patients through each individual EDSS using an annual cycle length and 20-year time horizon. During any cycle, patients may progress to a worse health state, remain in their current state, drop off therapy or die. Alongside these events, patients with RRMS may also develop SPMS. Costs and health utility scores are assigned to each state; as patients progress through the model they accrue costs and QALYs. Relative hazard ratios are applied to each transition rate to simulate the effect of disease-modifying therapies in delaying disease progression and preventing relapse. The use of disease-modifying therapies alters the trajectory of patients through the EDSS states and the number of relapses experienced, ultimately resulting in different profiles of costs and QALYs for each treatment strategy.

Owing to the time-dependence of the probabilities of transiting between health states, a single run of the model required around 7 seconds, meaning that probabilistic sensitivity analysis using the model was highly computationally expensive. In fact, if 10,000 samples were used for both the inner and outer levels, then calculating the partial EVPI for just 1 of the models 128 uncertain parameters using the 2-level sampling algorithm would require around 22.2 years computation time. Using

10,000 random samples we estimated the overall EVPI for the model to be £8,855 per patient. One-way sensitivity analyses suggested several model parameters as important and we chose to focus our analysis of the bias and confidence intervals for EVPI estimates on the parameter: mean cost associated with the EDSS 9.5 health state.

5 Results

5.1 Results using Case Study 1

The true EVPI for the parameter of interest X_5 is estimated at around \$199, which is 15.1 when indexed against overall EVPI (\$1319=100). The results of running the algorithm to predict bias and confidence intervals are shown in Table 3.

The predicted bias in the Monte Carlo estimate for a given inner sample size is compared with the actual bias exhibited in Figure 1. The order of magnitude of the predicted expected bias is very similar to that mean bias actually exhibited using 1000 repeated runs of the two level partial EVPI. The absolute scale of the bias is quite high for very small numbers of inner samples. For example, with just 10 inner runs used, the bias in this case study is approximately 15.5 when indexed to the overall EVPI = 100, which would more than double our estimate for the indexed partial EVPI for parameter X_5. When larger numbers of inner runs are used, the bias fall substantially with an expected indexed bias of just 0.25 when using $J = 500$ inner iterations. This suggests that on average, if we used $J = 500$ iterations we

	$J = 10$	$J = 100$	$J = 500$	$J = 1000$	$J = 5000$	$J = 10000$
Bias (indep't of K)	15.61	1.58	0.25	0.17	0.05	0.02
95% CI						
$K = 10$	±43.9	27.7	25.5	25.1	24.9	24.8
$K = 100$	±13.9	8.8	8.1	7.9	7.9	7.9
$K = 500$	±6.2	3.9	3.6	3.6	3.5	3.5
$K = 1000$	±4.4	2.8	2.5	2.5	2.5	2.5
$K = 5000$	±2.0	1.2	1.1	1.1	1.1	1.1
$K = 10000$	±1.4	0.9	0.8	0.8	0.8	0.8

Table 3: Predicted bias and 95% CI for Monte Carlo partial EVPI estimate in case study 1 using our proposed algorithm

would estimate the indexed partial EVPI for parameter X_5 at 15.35 rather than 15.1. The near equivalence of the predicted versus actuals for the different inner numbers of iterations tested suggests that our algorithm provides a robust estimate of the expected bias due to small numbers of iterations, at least in this case study.

Figure 3 shows that the predicted confidence intervals for the partial EVPI estimates are similar but slightly wider than the actual confidence intervals exhibited when using 1000 repeated computations of $EVPI(X_5)$. The estimated widths are of sufficient accuracy to be informative, and as discussed earlier, K can be refined once we have actually chosen J and obtained a partial EVPI estimate.

With the smallest numbers tested, ($K = J = 10$), the predicted 95% confidence

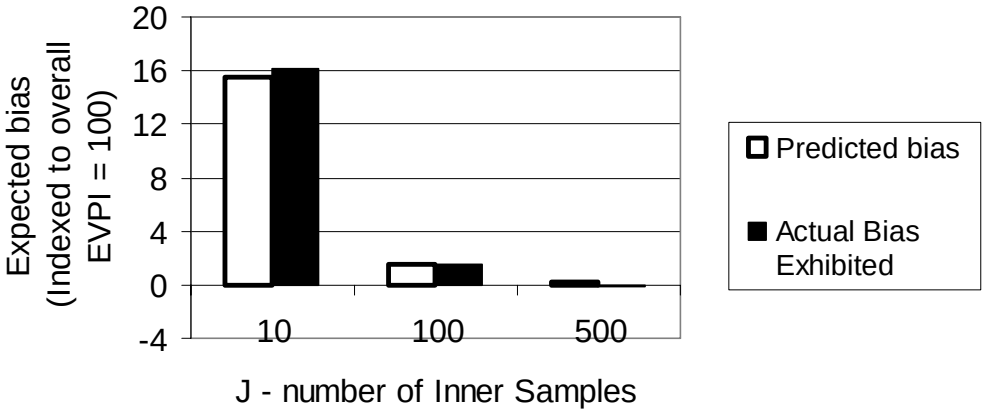

Figure 2: Comparison of Bias Predicted by Proposed Algorithm versus Exhibited bias in Case Study 1.

interval for the indexed partial EVPI is ± 43.8 i.e. an estimate could be expected to be biased by 15.5 indexed points as discussed earlier *and* vary anywhere between (-13 to +74) when indexed to overall EVPI = 100. Clearly using such small numbers to produce one estimate of $EVPI(X_5)$ would give almost meaningless results. As larger numbers of iterations are used, the predicted confidence intervals for $EVPI(X_5)$ reduce. When $K = J = 500$ is used, the predicted 95% confidence interval is 15.35 ± 3.6. It may be that even higher numbers of iterations are required if this level of accuracy is not enough for analysts or decision makers' needs. Again, the near equivalence of the predicted versus actuals for the different numbers of outer and inner iterations tested suggests that our algorithm provides a robust estimate of the order of magnitude of 95% confidence intervals for partial EVPI estimates, at least in this case study.

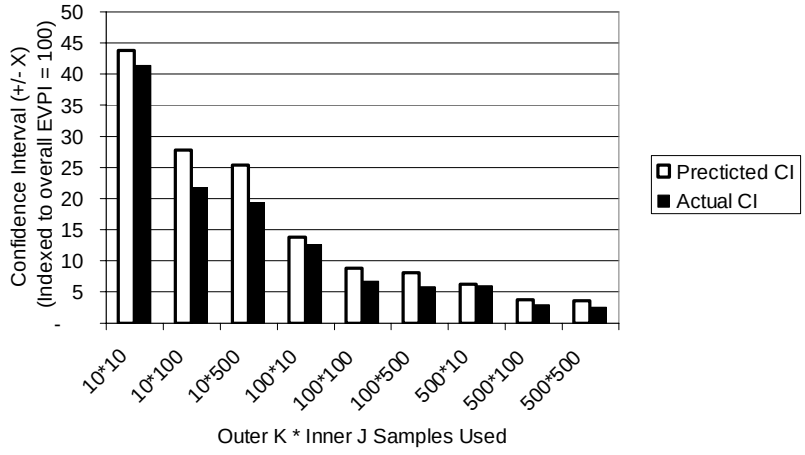

Figure 3: comparison of Confidence Intervals Predicted by Proposed Algorithm versus Exhibited Confidence Intervals in Case Study 1.

5.2 Results using Case Study 2

The true EVPI for the parameters of interest X_{20} to X_{31} is estimated at \$1677, which is 82 when indexed against overall EVPI (\$2045=100), a much larger proportion of the overall EVPI than the case study 1 example. The results of running the algorithm to predict bias and confidence intervals are shown in Table 4.

The predicted bias in the Monte Carlo estimate for a given inner sample size is compared with the actual bias exhibited in Figure 4. Again, the order of magnitude of the predicted expected bias is very similar to that mean bias actually exhibited using 1000 repeated runs of the two level partial EVPI.

Figure 5 shows that the predicted confidence intervals for the partial EVPI estimates are narrower than the actual confidence intervals exhibited when using 1000

	$J = 10$	$J = 100$	$J = 500$	$J = 1000$	$J = 5000$	$J = 10000$
Bias (indep't of K)	2.62	0.27	0.00	-0.02	0.00	-0.00
95% CI						
$K = 10$	±97.9	94.6	94.3	94.2	94.2	94.2
$K = 100$	±31.0	29.9	29.8	29.8	29.8	29.8
$K = 500$	±13.8	13.4	13.3	13.3	13.3	13.3
$K = 1000$	±9.8	9.5	9.4	9.4	9.4	9.4
$K = 5000$	±4.4.	4.2	4.2	4.2	4.2	4.2
$K = 10000$	±3.1	3.0	3.0	3.0	3.0	3.0

Table 4: Predicted bias and 95% CI for Monte Carlo partial EVPI estimate in case study 2 using our proposed algorithm

repeated computations of $EVPI(X_{20:31})$, but are still reasonably good estimates in this context, and satisfactorily show the relationship between the choice of K and the confidence interval width. We re-tested the algorithm using 101 outer samples in step 1 rather than 21. The results, plotted in figure 6 show considerably closer alignment between actual and predicted confidence interval widths, as we would expect.

5.3 Results using Case Study 3

For case study 3, the predicted bias in EVPI estimates produced by our algorithm for different numbers of inner samples J are shown in Table 3. The analysis suggests

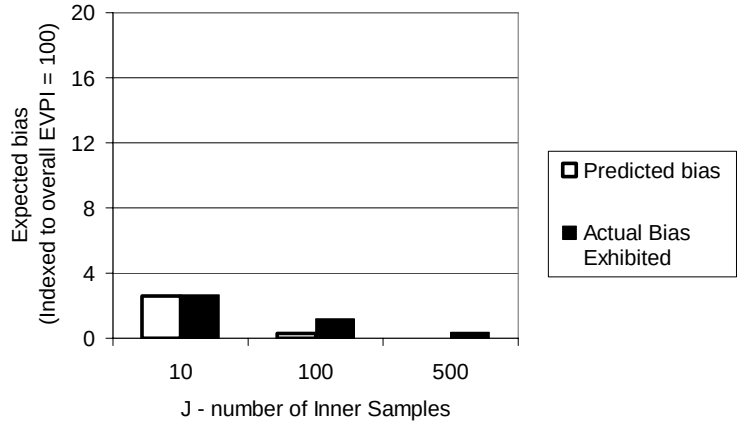

Figure 4: Comparison of Bias Predicted by Proposed Algorithm versus Exhibited bias in Case Study 2.

that an inner sample of 1,000 would still produce an appreciable bias (2.42 on the indexed scale).

In this case study, we tested different versions of our algorithm, extending the number of inner iterations used within the algorithm from our originally specified $J = 30$ up to 40, 50 and 60. We found some small variations but no appreciable differences in assessing the order of magnitude of the predicted bias when using greater than 30 inner iterations in our algorithm.

Table 6 shows the estimated the width of a 95% confidence interval for the partial EVPI for various outer sample sizes, given an inner sample size of $J = 1000$. We can see that the confidence intervals are fairly wide here, indicating a relatively large outer sample size will also be needed, certainly in excess of $K = 1000$.

The results of our case study 2 analyses, on the basis of 630 model evaluations, led us to establish that a likely minimum sample size for an accurate 2-level Monte

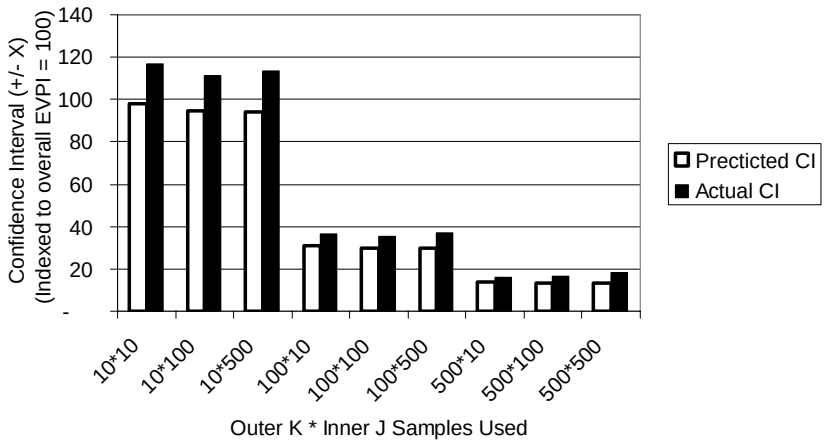

Figure 5: Comparison of confidence interval widths predicted by the proposed algorithm versus exhibited confidence interval widths in Case Study 2.

Carlo estimate will be of the order of 1 million model runs. Partial EVPI estimation using the Monte Carlo approach was clearly infeasible because this would require 81 days computation time. At this point we moved on to utilise the more efficient Gaussian Process model to approximate partial EVPI, the details of which are given elsewhere (Tappenden et al., 2004).

6 Discussion

We have presented an algorithm for predicting the expected bias and confidence interval widths for a Monte Carlo estimate of partial EVPI. Testing of the algorithm against exhibited bias and confidence intervals in our first and second case studies show that it provides predictions with the correct order of magnitude. Our third case study provides an example where infeasibly long computation times would be re-

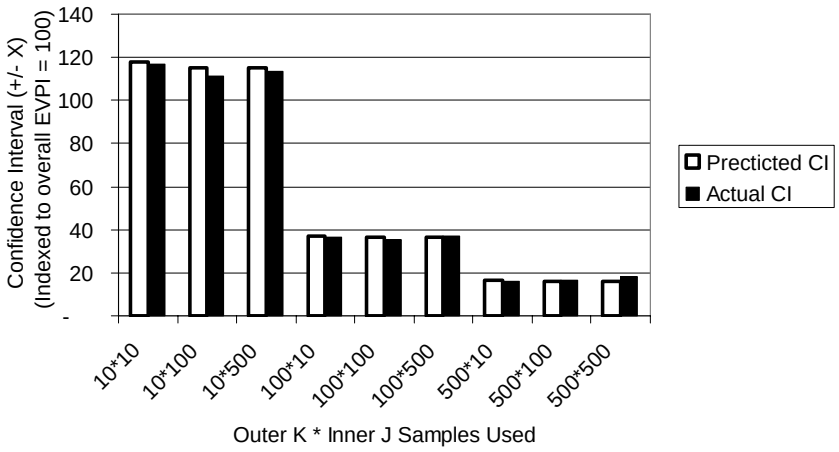

Figure 6: Comparison of confidence interval widths predicted by the algorithm with 101 outer samples at step 1 versus exhibited confidence intervals in Case Study 2.

quired for accurate EVPI estimates and a Gaussian process meta-model was required to emulate the original model and produce much quicker model runs (Tappenden et al., 2004).

The algorithm can be applied generally to any decision model, and is not limited to economic evaluation of health technologies. It can be applied to patient simulation models (Brennan et al., 2006), where there is the additional complication of how many patient simulations per parameter set to use (O'Hagan et al., 2007). It is also applicable to models with any characterisation of uncertainty and is not limited to continuous or parametric distributions.

The algorithm is very quick to compute. The time taken to compute predicted biases and confidence intervals for the 36 options of inner J and outer K samples in table 2 for case study 1 was just over 8 seconds. This compares with several days computation to produce the 36,000 runs of actual two level partial EVPIs with

Inner Samples	Predicted Bias	Indexed (overall EVPI=100)
$J = 500$	£ 456	5.25
$J = 1000$	£ 214	2.42
$J = 5000$	£ 45	0.51
$J = 10000$	£ 20	0.22

Table 5: The predicted size of a bias in a Monte Carlo partial EVPI estimate for parameter cost associated with health state EDSS 9.5, in case study 2.

Outer Samples K	K=100	K=1,000	K=10,000	K=100,000
Absolute Value	£ 11,729	£ 3,709	£ 1173	£ 371
Indexed to Overall EVPI	132.45	41.88	13.25	4.19

Table 6: Estimated Width of 95% confidence interval for a Monte Carlo partial EVPI estimate for parameter cost associated with health state EDSS 9.5, for an inner sample size $J = 1000$ and outer sample size K.

which to compare the predictions in Figures 2 and 3.

A limitation is that this procedure will be most useful for models that can run fairly quickly (i.e. a few seconds per model run). For very computationally expensive models, such as patient simulation models that can take an hour per run, we believe a Gaussian meta-model approach is currently the only viable option for computing EVPI.

One implication of considering the uncertainty in EVPI estimates explicitly is

that decision makers and analysts may need to define some limit on how accurately they wish to know the true EVPI for parameters of interest. Both case studies suggest that relatively small numbers of inner and outer loops, such as $K = J = 500$ and a total of 250,000 model runs, provide a reasonable estimate of the order of magnitude of the partial EVPI. The diminishing returns of higher numbers of inner and outer iterations for EVPI accuracy are an unavoidable fact of the Monte Carlo estimation process. In case study 1 moving from $K = J = 500$ to $K = J = 1000$ (an additional 750,000 models runs taking 4 times longer) reduces the 95% CI from ± 3.6 to ± 2.5 on the indexed scale. At a further extreme, moving from $K = J = 5,000$ to $K = J = 10,000$, an additional 75 million model runs reduces the 95% CI by just 0.3, from ± 1.1 to ± 0.8.

As partial EVPI calculations become more common in practice, we have a need to be as efficient as possible in producing estimates (Claxton et al., 2004). We have tested the algorithm on only two case studies here and further research would be useful to test the performance of the algorithm in other models. The field of value of information analysis is widening to consider expected value of sample information (EVSI) (Ades et al., 2004), (Brennan and Kharroubi, 2007a) and the expected value of perfect implementation (EVPImp) (Fenwick and Claxton, 2005). In many of these analyses, the same issue of inner and outer sampling exists and it would be useful to adapt our algorithm to these contexts and investigate its accuracy and usefulness.

In conclusion, this novel algorithm is easily and generally applied to compute predicted bias and confidence intervals for Monte Carlo based estimates of partial

EVPI in decision models. The algorithm is a relatively simple tool, using standard results concerning uncertainty in Monte Carlo estimates, but extending them into the context of nested expectations with intervening maxima. Our judgement is that it should provide robust estimates in all kinds of decision models but further testing of this may be useful. The algorithm is particularly useful when relatively small numbers of inner and outer iterations are planned and is recommended for use when reporting all Monte Carlo based partial EVPI calculations.

7 Acknowledgement

We thank two referees for their helpful comments on this paper.

References

Ades, A. E., Lu, G. and Claxton, K. (2004). Expected value of sample information calculations in medical decision modeling., *Medical Decision Making*, **24**: 207–227.

Brennan, A. and Akehurst, R. (2000). Modelling in health economic evaluation. What is its place? What is its value?, *Pharmacoeconomics*, **17**: 445–459.

Brennan, A., Chick, S. and Davies, R. (2006). A taxonomy of model structures for economic evaluation of health technologies, *Health Economics*, **15**: 1295–1310.

Brennan, A., Kharroubi, S., Chilcott, J. and O'Hagan, A. (2002). A

two level Monte Carlo approach to calculating expected value of perfect information:- Resolution of the uncertainty in methods. Available at http://www.shef.ac.uk/content/1/c6/02/96/05/Brennan.doc, Discussion paper, University of Sheffield.

Brennan, A. and Kharroubi, S. A. (2007a). Efficient computation of partial expected value of sample information using bayesian approximation, *Journal of Health Economics*, **26**: 122–148.

Brennan, A. and Kharroubi, S. A. (2007b). Expected value of sample information for weibull survival data, *Health Economics*, **16**: 1205–1225.

Brennan, A., Kharroubi, S. A., O'Hagan, A. and Chilcott, J. (2007). Calculating partial expected value of perfect information via monte-carlo sampling algorithms, *Medical Decision Making*, **27**: 448–470.

Briggs, A. and Gray, A. (1999). Handling uncertainty when performing economic evaluation of healthcare interventions., *Health Technology Assessment*, **3**.

Chilcott, J., Brennan, A., Booth, A., Karnon, J. and Tappenden, P. (2003a). The role of modelling in prioritising and planning clinical trials., *Health Technology Assessment*, **7**.

Chilcott, J., McCabe, C., Tappenden, P., O'Hagan, A., Cooper, N. J., Abrams, K. and Claxton, K. (2003b). Modelling the cost effectiveness of interferon beta and glatiramer acetate in the management of multiple sclerosis., *BMJ*, **326**: 522–528.

Claxton, K. (1999). Bayesian approaches to the value of information: implications for the regulation of new health care technologies, *Health Economics*, **8**: 269–274.

Claxton, K., Ginnelly, L., Sculpher, M., Philips, Z. and Palmer, S. (2004). A pilot study on the use of decision theory and value of information analysis as part of the NHS health technology assessment programme, *Health Technology Assessment*, **8**.

Claxton, K. and Posnett, J. (1996). An economic approach to clinical trial design and research priority-setting., *Health Economics*, **5**: 513–524.

Claxton, K. and Thompson, K. (2001). A dynamic programming approach to efficient clinical trial design., *Journal of Health Economics*, **20**: 797–822.

Felli, J. and Hazen, G. (2003). Correction: Sensitivity analysis and the expected value of perfect information, *Medical Decision Making*, **23**: 97.

Felli, J. C. and Hazen, G. B. (1998). Sensitivity analysis and the expected value of perfect information, *Medical Decision Making*, **18**: 95–109.

Fenwick, E. and Claxton, K and, S. M. (2005). The value of implementation and the value of information: Combined and uneven development, *CHE Discussion Paper 5*.

Koerkamp, B., Hunink, M., Stijnen, T. and Weinstein, M. (2006). Identifying key parameters in cost-effectiveness analysis using value of information: a comparison of methods, *Health Economics*, **15**: 383–392.

Meltzer, D. (2001). Addressing uncertainty in medical cost effectiveness analysis. implications of expected utility maximization for methods to perform sensitivity analysis and the use of cost-effectiveness analysis to set priorities for medical research., *Journal of Health Economics*, **20**: 109–129.

Oakley, J. and O'Hagan, A. (2004). Probabilistic sensitivity of complex models: a Bayesian approach, *J. Roy. Statist. Soc. Ser. B*, **66**: 751–769.

Oakley, J. E. (2009). Decision-theoretic sensitivity analysis for complex computer models, *Technometrics*, **51**: 121–129.

O'Hagan, A., Stevenson, M. and Madan, J. (2007). Monte Carlo probabilistic sensitivity analysis for patient level simulation models: efficient estimation of mean and variance using ANOVA, *Health Economics*, **16**: 1009–1024.

R Development Core Team (2006). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.

Raiffa, H. (1968). *Decision Analysis: Introductory Lectures on Choice Under Uncertainty*, Reading, Mass.: Addison-Wesley.

Stevenson, M., Oakley, J. and Chilcott, J. (2004). Gaussian process modelling in conjunction with individual patient simulation modelling: A case study describing the calculation of cost-effectiveness ratios for the treatment of osteoporosis, *Medical Decision Making*, **24**: 89–100.

Tappenden, P., Chilcott, J., Eggington, S., Oakley, J. and McCabe, C. (2004). Methods for expected value of information analysis in complex health economic models: developments on the health economics of interferon- and glatiramer acetate for multiple sclerosis., *Health Technology Assessment*, **8**.

Van Hout, B., Al, M., Gordon, G. and Rutten, F. (1994). Costs, effects and c/e-ratios alongside a clinical trial., *Health Economics*, **3**: 309–319.

Yokota, F. and Thompson, K. (2004). Value of information literature analysis (VOILA): A review of applications in health risk management., *Medical Decision Making*, **24**: 287–298.