Expected Value of Sample Information Calculations in Medical Decision Modeling

A. E. Ades, PhD, G. Lu, K. Claxton, PhD

There has been an increasing interest in using expected value of information (EVI) theory in medical decision making, to identify the need for further research to reduce uncertainty in decision and as a tool for sensitivity analysis. Expected value of sample information (EVSI) has been proposed for determination of optimum sample size and allocation rates in randomized clinical trials. This article derives simple Monte Carlo, or nested Monte Carlo, methods that extend the use of EVSI calculations to medical decision applications with multiple sources of uncertainty, with particular attention to the form in which epidemiological data and research findings are structured. In particular, information on key decision parameters such as treatment efficacy are invariably available on measures of relative efficacy such as risk differences or odds ratios, but not on model parameters themselves. In addition, estimates of model parameters and of relative effect measures in the literature may be heterogeneous, reflecting additional sources of variation besides statistical sampling error. The authors describe Monte Carlo procedures for calculating EVSI for probability, rate, or continuous variable parameters in multiparameter decision models and approximate methods for relative measures such as risk differences, odds ratios, risk ratios, and hazard ratios. Where prior evidence is based on a random effects meta-analysis, the authors describe different ESVI calculations, one relevant for decisions concerning a specific patient group and the other for decisions concerning the entire population of patient groups. They also consider EVSI methods for new studies intended to update information on both baseline treatment efficacy and the relative efficacy of 2 treatments. Although there are restrictions regarding models with prior correlation between parameters, these methods can be applied to the majority of probabilistic decision models. Illustrative worked examples of EVSI calculations are given in an appendix. ***Key words****: expected value of sample information; epidemiology; odds ratios; random effects.* ***(Med Decis Making 2004;24:207–227)***

A societal decision maker confronted with a choice of alternative health care interventions should choose the intervention with the greatest expected net benefit if he or she wishes to maximize heath outcome subject to a budget constraint.[1,2] However, in the presence of uncertainty about the net benefits associated with the alternative treatments, there is a finite probability that the optimal decision is wrong. The expected opportunity loss associated with this uncertainty can be quantified and is known as the expected value of perfect information (EVPI) because it is the amount the decision maker should be willing to pay to eliminate all uncertainty in the decision. These concepts were developed within Bayesian statistical decision theory.[3,4]

Where there are multiple sources of uncertainty, it is possible to calculate the expected value of partial perfect information (EVPPI), the EVPI on a subset of parameters. Several authors both in the risk analysis literature[5–7] and in medical decision literature[8–10] have highlighted the use of these methods in research prioritization and in sensitivity analysis.

The EVPI and EVPPI set an upper limit on the societal returns to further research. They provide a "hurdle" since further investigation will be potentially worthwhile only if the EVPI exceeds the cost of further research.[11] However, complete elimination of uncertainty can be achieved only by an infinitely large sam-

Received 29 November 2002 from the Medical Research Council Health Services Research Collaboration, Bristol, United Kingdom (AEA, GL), and the Department of Economics and Related Studies, University of York, United Kingdom (KC). The authors would like to thank the reviewers of *Medical Decision Making* for their careful reading and valuable suggestions and to acknowledge the contributions of Professor Tony O'Hagan, who commented on an earlier draft, and other colleagues during the Centre For Bayesian Statistics in Health Economics Focus Fortnight on Expected Value of Information, Department of Probability and Statistics, University of Sheffield, 13–24 May 2002. Any errors are the authors' responsibility. Revision accepted for publication 12 November 2003.

Address correspondence and reprint requests to A. E. Ades, PhD, MRC Health Services Research Collaboration, Canynge Hall, Whiteladies Road, Bristol BS8 2PR, England; e-mail: t.ades@bristol.ac.uk.

DOI: 10.1177/0272989X04263162

ple. The practical task is therefore to calculate the expected value of sample information (EVSI) and to set this against the costs of obtaining the sample. The difference between the reduction in the expected loss due to sample information and the costs of obtaining the sample is the expected net benefit of sampling and represents the societal return to proposed research.[3,4] The EVSI calculation is therefore part of the process of finding an optimum sample size for a future study. EVSI studies can be found particularly in the environmental engineering and health risk assessment literature.[12–14] In the medical decision context, Claxton has applied EVSI to the optimal size and treatment allocation in randomized clinical trials (RCTs).[2,11,15]

Decision theoretic EVI analysis provides an explicit, coherent, and flexible framework with many advantages. It separates the decision to adopt a health care technology given current evidence from the decision on whether more research is required to inform this choice. This contrasts with the traditional inference-led approach in which the adoption decision is based on arbitrary significance tests or confidence ranges, research design is based on arbitrary power calculations, and there is no guidance on which model parameters are most deserving of further data. Furthermore, the explicit valuation of both the benefits and the costs of obtaining more information produces an optimal study size, rather than a perpetual plea that "more research is needed"[16] or that more "large"[17] or even "very large" studies are needed.[18] In contrast, EVI theory not only provides a framework for efficient research design but also guarantees that research expenditure is commensurate with service provision,[11] as well as offering a coherent framework for regulatory authorities.[19]

EVI concepts are now beginning to attract attention in the epidemiological literature,[16] but despite their power and coherence, practical full-scale applications of EVSI methods are difficult to find. Besides the complexity of EVI calculations and the difficulty of the traditional notation used in textbooks, there are also conceptual obstacles. Based as it is on Bayesian decision theory, EVI requires probabilistic decision modeling,[20–22] which is only recently becoming accepted.[23] This methodology requires investigators to be completely specific about parameter uncertainty and thereby raises very searching questions about the origin of uncertainty and variation in epidemiological data. These questions can readily be sidestepped in deterministic sensitivity analysis, but probabilistic modeling requires that all sources of uncertainty and variation be appropriately represented and no more so than in EVI calculations. After all, if uncertainty is a result of the limitations of current data and is to be reduced by collecting more data, it follows that careful attention has to be paid toward how epidemiological data are structured, interpreted, and analyzed.

An additional impetus for this work is the increasing use of formal health economic assessment in decision models intended to inform national policy on the provision and reimbursement of specific heath care technologies.[24–28] Such analyses are intended primarily to inform the decision to adopt or reimburse a technology based on current evidence. However, these models are based on increasingly sophisticated and thorough use of systematic literature review and meta-analysis, and their output provides a natural source for consensus-based prior parameter distributions for input to EVSI analyses, which would directly address whether additional evidence is required to support a technology. This issue is, indeed, now actively considered by many regulatory and reimbursement authorities.[28]

This article begins with a brief introduction to EVI concepts to establish a conceptual framework and notation (section 1). We then develop EVSI calculations for probability, rate, and normally distributed variables (section 2). The implications of the "fixed" versus "random effects" distinction[29–31] are pursued in section 3. Section 4 shows how information on relative effect measures, such as odds ratios, can be mapped into information on the absolute effect measures, the latter being the more relevant in decision models. In discussion, we examine the limitations and the scope of the proposed methods, the approximations made on the way, and the wider optimization problem design of an optimal research portfolio. An appendix provides some worked examples of EVSI calculations.

1. FRAMEWORK FOR EXPECTED VALUE OF INFORMATION

1.1. Incremental Net Benefit

We adopt a net benefit approach to cost-effectiveness analysis.[1,11] Health gains are monetarized by multiplying quality adjusted life years (QALYs) by a cost-effectiveness threshold λ, the decision maker's presumed willingness to pay per additional QALY. This is gaining acceptance[32–35] not least because it avoids the difficulties associated with cost-effectiveness ratios.

We assume a decision model with unknown parameters θ, with a choice to be made between a fixed number of treatments $t = 1,2 \ldots T$. $B(t, \theta)$ is the net benefit of treatment t if the parameters take the value θ. The net benefit attaching to treatment t is thus

$$B(t, \theta) = \lambda U(t, \theta) - C(t, \theta), \quad (1)$$

where the functions C and U give the costs and the QALYs under treatment t.

1.2. Expected Value of Perfect Information (EVPI)

The optimal decision given current information is the decision that yields the highest expected net benefit:

$$\max_t E_\theta B(t, \theta).$$

The true values of θ are not known, but if they were known, we could maximize over t, $\max_t B(t, \theta)$, to obtain the value of an optimal decision at these known values of θ. As the θ are not known, the expected net benefit of a decision taken with perfect information is found by averaging this expression over the joint distribution of θ:

$$E_\theta \max_t B(t, \theta).$$

The EVPI is the difference between the expected value of a decision made with perfect information about all the uncertain parameters θ and a decision made now:

$$\text{EVPI} = E_\theta \max_t B(t, \theta) - \max_t E_\theta B(t, \theta). \quad (2)$$

1.3. Expected Value of Partial Perfect Information (EVPPI)

Suppose we were interested in the value of perfect information about a vector subset θ_I of the parameters θ. If we knew the values of θ_I, then the expected net benefit of a decision made now would be found by averaging over the uncertainty in θ that remains after knowing θ_I and then selecting the optimal treatment

$$\max_t E_{\theta|\theta_I} B(t, \theta).$$

But again, we do not know θ_I, so the expected value of a decision made with perfect information about θ_I is found by averaging over our current information about θ_I:

$$E_{\theta_I} \max_t E_{\theta|\theta_I} B(t, \theta).$$

The EVPPI is the difference between the expected value of a decision made with perfect information about θ_I and the current optimal decision:

$$EVPPI = E_{\theta_I} \max_t E_{\theta|\theta_I} B(t, \theta) - \max_t E_\theta B(t, \theta). \quad (3)$$

Consider the parameters of interest θ_I and the complement set θ_I^c. It has been noted[8,9] that if $B(t, \theta)$ is linear in θ_I^c and there are no correlations between the θ_I and the θ_I^c, then we can rewrite as $E_{\theta|\theta_I} B(t, \theta_I, \theta_I^c)$ as $B(t, \theta_I, E(\theta_I^c))$ and thus avoid the inner integration. The short cut it also possible when $B(t, \theta_I, \theta_I^c)$ is multilinear in the components of θ_I^c and there are no correlations between the components of θ_I^c or between the θ_I and the θ_I^c. This condition allows the net benefit functions to contain products of (independent) parameters. Therefore, many standard decision tree models, with branching path probabilities and independent parameters, are amenable to EVPPI calculations that require only a single Monte Carlo integration. In general, a model is multilinear in a parameter as long as it occurs only once on any path from the origin to a terminal node. However, a nested inner integration would be required for Markov models, which bring in higher order terms in branch probabilities.

1.4. Expected Value of Perfect Information at a Population Level

The expressions for EVPI and EVPPI refer to expected value of information per individual patient presenting at the decision point. These expressions need to be scaled up to a population level to reflect the effective EVI that is relevant to a decision about further research.[2,11] If Q_y patients (or patient episodes) enter this decision problem in year y, with an H year effective lifetime for the technology and a discount rate of α, then

$$\text{Population EVPI} = \text{EVPI} \cdot \Sigma_{y=1,2\ldots H} Q_y (1+\alpha)^{-y}.$$

Both the number of patients entering the decision problem each year and the effective lifetime of the technology may themselves be unknown parameters on which further information could be sought. It should be noted that with a positive discount rate, the population EVPI will be finite.

1.5. Expected Value of Sample Information (EVSI)

If EVPPI calculations suggest there are P parameters on which it may be cost-effective to obtain more information (i.e., the EVPPI exceeds the cost of conducting further research), the question of which information to obtain and how much could represent an extremely large decision problem (see section 1.6). Here, we assume that a particular study is being considered with a specific sample size vector n. The new study will pro-

vide sufficient statistics D relating to parameters of interest θ_I.

The development mirrors that of EVPPI. If we knew in advance what D would be, the expected value of a decision made after sufficient statistics D have been acquired is found by taking an expectation over the posterior distribution of the net benefit of each treatment t given the new data D, that is,

$$\max_t E_{\theta|D}\, B\,(t, \theta).$$

If we suppose that θ_I and θ_I^c are a priori independent and that the data D provide information only about the subvector θ_I and not the complement set θ_I^c, then this expression can further be written as

$$\max_t E_{\theta_I^c,(\theta_I|D)} B(t, \theta_I, \theta_I^c),$$

where the expectation is taken over the prior density of θ_I^c and the posterior density of θ_I given D.

As we do not yet know what D will be, we must average over the distribution of D:

$$E_D \max_t E_{\theta_I^c,(\theta_I|D)} B(t, \theta_I, \theta_I^c).$$

Finally, the distribution of D must be based on prior knowledge of θ_I. To clarify the nature of the computation, we can consider the expectation over D as an expectation over the predictive distribution of the new data D conditional on θ_I, averaged over the prior distribution of θ_I.

As before, the expected value of sample information is the difference between the expected value of a decision made after data D have been collected and the expected value of a decision made now:

$$E_D \max_t E_{\theta_I^c,(\theta_I|D)} B(t, \theta_I, \theta_I^c) - \max_t E_\theta B(t, \theta_I, \theta_I^c). \quad (4)$$

In parallel with EVPPI, this expression contains an inner integration within the maximization step. This takes the form of the Bayesian calculation of a posterior mean, after priors θ_I have been combined with new data D. This article restricts itself to situations in which the likelihood for the proposed data D is conjugate with priors, so that means or parameters for posterior distributions are available in closed form and to scalar priors with no correlations. These restrictions are strong but nevertheless allow for EVSI calculations via single or nested (see below) Monte Carlo simulations in a very wide class of decision models. Possible approaches for a still wider class of models are taken up in discussion.

The following algorithm shows how the calculation of EVSI can be implemented in a Monte Carlo decision-modeling framework.

General Algorithm for Calculation of EVSI

A1. For $i = 1,2 \ldots N$ simulations
- B1. Draw a sample $\theta_I^{(i)}$ from the prior distribution of θ_I.
- B2. Draw a sample $D^{(i)}$ from the distribution of the sufficient statistics $D \mid \theta_I^{(i)}$ arising from a new study of size n
- B3. Calculate posterior expected net benefits for each strategy t, using algorithms C1, C2, C3, and C4 (see below) as appropriate
- B4. Find the treatment t maximizing expected net benefit for simulation i and record the corresponding maximum value

A2. Find the average of the maximum expected net benefits, over the N simulations. This is the expected value of a decision based on sample information

A3. Subtract the expected value of a decision based on current information

---

The main focus of the article will be to set out steps B1, B2, and B3 for a wide range of data types and prior evidence structures. We start with the calculations C1, C2, C3, and C4 required in step B3 to produce the expected posterior net benefits over which the maximization must take place on each simulation,

$$E_{\theta_I^c,(\theta_I|D^{(i)})} B(t, \theta_I, \theta_I^c). \quad (5)$$

The choice of algorithm is determined by the linearity of the net benefit functions in θ_I and θ_I^c. Four algorithms to calculate this "inner" integration can be distinguished, all assuming prior independence of θ_I and θ_I^c.

Algorithms to Find the Expectation of Net Benefits after New Data Are Collected

C1. $B(t, \theta)$ is linear in θ_I and θ_I^c AND there are no correlations between any of the θ_I and θ_I^c parameters; OR $B(t, \theta)$ is multilinear and there are no correlations between any parameters: Plug in the prior means for the parameters θ_I^c and the posterior means for θ_I.

$$E_{\theta_I^c,(\theta_I \mid D^{(i)})} B(t, \theta_I, \theta_I^c) = B(t, E(\theta_I \mid D^{(i)}), E(\theta_I^c)).$$

For example, $B(t, \theta_1, \theta_2)$ is multilinear (but not linear) in its parameters when it contains products of independent parameters.

C2. $B(t, \theta)$ is linear in θ_I but nonlinear in θ_I^c: Carry out a nested Monte Carlo integration over θ_I^c, drawing from the prior distributions of the parameters θ_I^c and plugging in the posterior means for θ_I.

$$E_{\theta_I^c,(\theta_I \mid D^{(i)})} B(t, \theta_I, \theta_I^c) = E_{\theta_I^c} B(t, E(\theta_I \mid D^{(i)}), \theta_I^c)$$

C3. $B(t, \theta)$ is nonlinear in θ_I but linear in θ_I^c: Carry out a nested Monte Carlo integration over the posterior $\theta_I \mid D$ and plug in the prior means for θ_I^c.

$$E_{\theta_I^c,(\theta_I \mid D^{(i)})} B(t, \theta_I, \theta_I^c) = E_{\theta_I \mid D^{(i)}} B(t, \theta_I, E(\theta_I^c)).$$

C4. $B(t, \theta)$ is nonlinear in both θ_I and θ_I^c: Carry out a nested Monte Carlo integration, drawing from the posterior distributions of the parameters $\theta_I \mid D$ and from the prior distributions of θ_I^c.

Thus, for algorithms C1 and C2, it is necessary to calculate only the posterior mean of the parameters on which new data are to be collected. For algorithms C3 and C4, it is necessary to have the parameters of the posterior distribution in closed form so that samples can be drawn from them. The most important distinction in practice, however, is between algorithm C1, which allows the posterior net benefits to be found by plugging in prior means of θ_I^c and posterior means of θ_I for that simulation, and the other algorithms that require a nested inner Monte Carlo integration.

1.6. Cost of Sampling and the Expected Net Benefit of Sampling

EVSI calculations are invariably set within an optimization problem relating to sample size. If we consider a study of known design with fixed sample size n, aimed at providing information on 1 or more parameters, then the expected net benefit of a research portfolio is the difference between the population EVSI for n and the cost of the portfolio[3]:

$$\text{ENBS}(n) = \text{Population EVSI}(n) - \text{Cost}(n),$$

and the task is to find the n that maximizes ENBS(n).

In medical applications to date, calculations of ENBS(n) have mainly been discussed in cases in which the research portfolio consists of a single 2-arm RCT, involving treatment-efficacy parameters for each arm. Optimization for n, which is then a vector of 2 sample sizes to be allocated to different treatments, can then proceed by calculating ENBS for a grid of values of $\{n_1, n_2\}$.[2] As we see in the next section, follow-up time F may be a further design feature to be optimized along with (n).

Although distinct from the EVSI calculation, it is worth briefly outlining the cost calculations. These include both direct resource costs and opportunity costs. The direct costs are the fixed costs of further research, and if the research portfolio might include an RCT, the marginal reporting costs and any additional treatment costs for patients allocated to a nonstandard treatment arm. Three elements of the opportunity costs of sampling can be distinguished. First, those who participate in the trial cannot benefit from the information generated by the research. Increasing sample size therefore provides more information for future patients but also "uses up" those who would otherwise be able to benefit from the sample information. Similarly, additional sample or longer follow-up, which delays the trial report, will also reduce the number of future patients who can benefit from the sample information. Third, while the trial is being conducted, patients not enrolled may receive standard treatment even when this is not a priori optimal due to concerns about the irreversibility of switching to the new technology. These patients will forgo the expected additional net benefit associated with the a priori optimal treatment. Claxton provided an outline of such calculations.[2] A similar approach has also been applied to sequential design problems in RCTs.[36]

Returning to the optimization problem, it must be emphasized that for a model with uncertainty in a number of parameters, the optimization problem may be difficult to solve. With P parameters on which further information could be usefully obtained, in the sense that they pass the EVPPI "hurdle," the full research space could be a P-tuple of sample sizes $n = \{n_1, n_2 \ldots n_P\}$. Clearly, a grid search solution may not be feasible if the research space is large. This is a problem to which we return briefly in discussion. It is not the role of this article to propose solutions but only to note that the potential complexity of the optimization problem suggests that we should take advantage of any approximations that might reduce computing time for EVSI(n).

2. UNCERTAINTY IN SINGLE PARAMETERS

2.1. Introduction

This section rehearses the standard results of Bayesian theory,[37,38] as they would be applied to single parameters, or sets of single parameters, within our Monte Carlo implementation of EVSI calculations. We

take 3 common parameter types—path probabilities, hazard rates, and continuous variables—and then look at the questions raised by heterogeneity in prior information. As demonstrated in the general EVSI algorithm, all that is required is that we can specify the parameters of the prior distribution in closed form, generate new data from the prior, update the priors with the new data, and either identify the posterior mean (algorithms C1 and C2) or the parameters of the posterior distribution (algorithms C3 and C4). Below, we describe the general algorithm for EVSI calculations for probability, rate, and normally distributed parameters, when the prior information relates directly to the parameters, not to a function of the parameters.

In each case, we assume that the prior information is based on an observation of n_0 individuals, although this can be a notional sample size if the prior is based on expert opinion. In the case of a continuous, normally distributed variable, the prior variance (of an individual observation) is σ_0^2, which can either be known or unknown. The prior variance is based on ν_0 degrees of freedom. If the prior variance is based on the sample variance of n_0 previous observation, then $\nu_0 = n_0 - 1$, but again, the degrees of freedom could be notional. It is often convenient to refer to precision, which is 1/variance. Throughout the article, we use the symbol τ for the precision and parameterize the normal distribution in terms of mean and precision: normal(μ, τ). In each case, the decision maker is considering a new study of size n for a single parameter. For a probability parameter, the prior is based on having observed a successes out of n_0 trials. In the case of a rate parameter, the prior is based on n_0 events observed over a total Y_0 years at risk, and the proposed study will observe n patients with an average F years of follow-up. It should be emphasized that the algorithms below refer to scalar parameters, but they can also be applied when θ_I includes multiple scalar parameters, as long as the restrictions on correlation (section 1.5) are met.

To see how the standard results of Bayesian theory apply in a Monte Carlo simulation framework, we distinguish between the name given to a random variable θ and the value it takes in the ith simulation, which is given a superscripted index $\theta^{(i)}$.

Specific Algorithms for EVSI for Individual Parameters, Steps B1–B3

For each outer simulation i:

B1. Draw a sample $\theta_I^{(i)}$ from the prior distributions of the parameters θ_I on which more data are to be collected.

Probability, based on a events out of n_0:
$\theta_I^{(i)} \sim \text{Beta}(a, n_0 - a)$

Rate, based on n_0 events in Y_0 total years at risk:
$\theta_I^{(i)} \sim \text{Gamma}(n_0, Y_0)$

Normal variable, known variance $\sigma_0^2 = 1/\tau_0$:
$\theta_I^{(i)} \sim \text{Normal}(\mu_0, n_0\tau_0)$

Normal variable, unknown variance. First draw a sample from the prior distribution of the population precision (= 1/variance), then draw a sample from the prior distribution of the mean given that precision:

$$\tau_0^{(i)} \sim \text{Gamma}(\nu_0/2, \nu_0\sigma_0^2/2)$$
$$\theta_I^{(i)} \sim \text{Normal}(\mu_0, n_0\tau^{(i)})$$

B2. Draw a sample from the predictive distribution of the sufficient statistics $D^{(i)}$ arising from a new study of size n, given the current $\theta_I^{(i)}$

Binomial numerator: $r_D^{(i)} \sim \text{Binomial}(\theta_I^{(i)}, n)$

Poisson event count: $e_D^{(i)} \sim \text{Poisson}(\theta_I^{(i)} nF)$

Sample mean, known variance:
$\mu_D^{(i)} \sim \text{Normal}(\theta_I^{(i)}, n\tau_0)$

Sample mean, unknown variance. First draw a sample $\tau_D^{(i)}$ from the predictive distribution of the population precision, then draw a sample $\mu_D^{(i)}$ from the predictive distribution of the sample mean:

$$\tau_D^{(i)} \sim \text{Gamma}((n-1)/2, (n-1)/(2\tau^{(i)}))$$
$$\mu_D^{(i)} \sim \text{Normal}(\theta_I^{(i)}, n\tau_D^{(i)})$$

B3. Posterior expectation $E(\theta_I \mid D^{(i)})$, for use in algorithms C1, C2:

Probability: $E(\theta_I \mid D^{(i)}) = (a + r_D^{(i)})/(n_0 + n)$

Rate: $E(\theta_I \mid D^{(i)}) = (n_0 + e_D^{(i)})/(Y_0 + nF)$

Normal variable, known variance:
$E(\theta_I \mid D^{(i)}) = (\mu_0 n_0 + \mu_D^{(i)} n)/(n_0 + n) = \mu_{post\text{-}kn}^{(i)}$

Normal variable, unknown variance:

$$E(\theta_I \mid D^{(i)}) = \frac{(\mu_0 n_0 \tau_0 + \mu_D^{(i)} n\tau_D^{(i)})}{(n_0\tau_0 + n\tau_D^{(i)})} = \mu_{post\text{-}ukn}^{(i)}$$

Posterior distributions, for use in C3, C4:

Probability: $\theta_I \mid D^{(i)} \sim \text{Beta}(a + r_D^{(i)}, n_0 + n - a - r_D^{(i)})$

Rate: $\theta_I \mid D^{(i)} \sim \text{Gamma}(n_0 + e_D^{(i)}, Y_0 + nF)$

Normal variable, known variance:
$\theta_I \mid D^{(i)} \sim \text{Normal}(\mu_{post\text{-}kn}^{(i)}, \tau_0(n_0 + n))$

Normal variable, unknown variance:

$$\theta_I \mid D^{(i)} \sim t_{\nu_0 + n}(\mu_{post-unk}^{(i)}, \tau_{post-unk}^{(i)})$$

$$\tau_{post-ukn}^{(i)} = (n + n_0)(\nu_0 + n)/[\nu_0/\tau_0 + n/\tau_D^{(i)} + n_0 n(\mu_0 - \mu_D^{(i)})^2/(n_0 + n)]$$

Table 1 Expectation $E(\theta)$ of a Continuous Variable θ, Which Is Normally Distributed after Transformation

Transformation $G(\theta)$	$E(\theta)$
Natural logarithm: $\log(\theta) \sim \text{Normal}(\mu, \tau)$, $\theta > 0$	$\exp(\mu + 1/(2\tau))$
Power transform*: $\theta^a \sim \text{Normal}(\mu, \tau)$, $\theta > 0$	$\mu^b[1 + b(b-1)/(2\tau\mu^2)]$, $b = 1/a$
Logistic*: $\log(\theta/(1-\theta)) \sim \text{Normal}(\mu, \tau)$, $0 < \theta < 1$	$h + h(1-h)(1-2h)/(2\tau)$, $h = e^{\mu}/(1+e^{\mu})$

Note: $G(\theta) \sim \text{Normal}(\mu, \tau)$. τ is precision (= 1/variance).
*Approximation by 2nd-order Taylor series expansion.[68]

We refer above to the posterior distribution because in each simulation, a posterior mean is created from the new data $D^{(j)}$ and the prior. However, this is ahead of any new data being collected. Over the N simulations, the steps B1–B3 are in fact building the prior distribution of the posterior mean, also known as the preposterior distribution, although we do not sample from this distribution explicitly. It is instructive to consider the variance of the preposterior distribution and its relation to sample size, as its role in the EVSI calculation is the same as the role of the prior in EVPI. Taking the case of a normally distributed variable with known variance as an example, it can be shown[3] that the variance of the preposterior distribution is

$$\frac{\sigma_0^2}{n_0}\left(\frac{n}{n+n_0}\right).$$

With a small sample size n, the new data, when we have it, will add very little, so we know in advance that our posterior mean will be close to our prior mean, which is already known. Consequently, the variance of the preposterior becomes lower as n decreases. As the sample size increases, the variance of a sample mean (σ_0^2/n) of course decreases, but we become increasingly uncertain about where the posterior mean will be. This will depend more and more on the sample mean, and uncertainty about the sample mean converges toward the uncertainty in the prior mean (σ_0^2/n_0), which is, after all, what is generating the sample. Thus, with an increasing sample size, the variance in the preposterior approaches the variance in the prior, until EVSI with an infinite sample is equal to EVPI.

2.2. Transformations of Data or Parameters

Continuous variables may be transformed to a log or other scale $\xi = G(\theta_I)$ to stabilize the variance and produce normality. Then the prior, the sufficient statistics for the new sample, and the formation of a posterior all take place on the transformed scale using the procedures for normally distributed data outlined above. Then, at step B3, we can follow algorithms C3 or C4 by drawing a value $\xi \mid D^{(j)}$ from the posterior distribution of the variable on the transformed scale and then back transforming to obtain a value of $\theta_I \mid D^{(j)}$ to enter into the net benefit function in an inner Monte Carlo integration. Alternatively, if the net benefit is linear in θ_I, then it may be helpful to use expressions for $E(\theta_I \mid D^{(j)})$ that can be obtained in closed form (see Table 1).

A similar situation obtains when the prior information and the new sample will provide information on a basic variable that is related to net benefit via a mathematical function. For example, if rate parameters are expressed as instantaneous hazards, cumulative discounted net benefits can readily be derived directly using stochastic tree models.[39–42] But this requires that the model be run over a lifetime and that hazard rates remain constant. As a result, hazard rates θ are frequently transformed to a discrete time form in which $\xi = G(\theta) = 1 - e^{-\theta X}$ is the probability of moving from 1 state to another in time X.[43,44] The procedure here would therefore be to adopt a nested Monte Carlo algorithm: The prior distribution would be updated on the scale of the rate parameter, then samples would be drawn from its posterior distribution and transformed into discrete time probabilities, to obtain the posterior mean net benefits via Monte Carlo integration. Care must be taken to convert continuous time Markov rates in multiple state models into discrete time Markov transition probability matrixes.[45]

3. HETEROGENEITY IN PRIOR OBSERVATIONS

3.1. Random Effects Models

If there are M independent sources of information on a parameter, these can be combined into a suitably weighted pooled estimate that can serve as a basis for a beta, gamma, or normal prior. This assumes that there is a single fixed effect of which each prior study gives an estimate. But the existence of more than 1 source of data raises the question of whether the estimates are

consistent. At the opposing extreme, one might regard all the M estimates as completely unconnected. An intermediate position, the random effects model is that the M study-specific estimates are drawn from a common distribution. (It should be emphasized that although a statistical significance test of the null hypothesis of no heterogeneity can provide strong evidence that there is heterogeneity, it is unlikely to produce convincing evidence that heterogeneity can be ruled out.) Fixed versus random effects models are usually discussed in the context of measures of relative efficacy,[46] and this is taken up in section 4. However, the same issues are relevant for any measure, and we now illustrate an approximate approach for a probability parameter.

First, we represent the available evidence D_0 from M previous studies statistically in terms of a random effects model. We assume that the logits of the probabilities in these studies, rather than the probabilities themselves, are from a common normal distribution. (One rationale for this rescaling is that all values for probability parameters are forced to be in the interval 0 to 1). Each study i, $i = 1,2 \ldots M$, has produced estimates y_i of the $\xi_i = \text{logit}(\theta_i)$, where θ_i is the "true" probability estimated by study i. Each y_i has a known variance σ_i^2 determined by the study numerator r_i and denominator n_i and based on a normal approximation. The data structure is

$$y_i = \log(r_i) - \log(n_i - r_i)$$

$$\sigma_i^2 = 1/r_i + 1/(n_i - r_i).$$

We assume a 2-level Gaussian model, in which the ξ_i are drawn from a common normal distribution representing the population of studies and the observed logits are also normally distributed about their expected values.

$$y_i \mid \xi_i \sim \text{Normal}(\xi_i, \sigma_i^2),$$

$$\xi_i \mid \mu \sim \text{Normal}(\mu, \sigma_\xi^2).$$

A range of methods is available to estimate such random effects models from the available data[46,47] and specifically to generate estimates of the population variance σ_ξ^2. In what follows, we adopt an approximate solution in which the value of $\sigma_\xi^2 = 1/\tau_\xi$ based on the original data D_0 is assumed fixed and known and will not be updated explicitly in the data-gathering exercise. Given such an estimate, we obtain estimates of the population mean μ, its precision τ_μ, and the ξ_i and their precisions τ_i, as follows:

$$\mu = \Sigma_i w_i y_i / \Sigma_i w_i,\ w_i = 1/(\sigma_i^2 + \sigma_\xi^2),\ \tau_\mu = 1/\Sigma_i w_i$$

$$\xi_i = (y_i/\sigma_i^2 + \mu/\sigma_\xi^2)/(1/\sigma_i^2 + 1/\sigma_\xi^2),\ \tau_i = 1/\sigma_i^2 + 1/\sigma_\xi^2.$$

The critical feature of the random effects model can be seen in the estimates ξ_i, which are "shrunk" toward μ by an amount that depends on distance from the mean and sample size. Their decreased variance, compared to the empirical variances σ_i^2, reflects the way the ξ_i borrow strength from their neighbors, through being drawn from a common underlying distribution.

The estimates from this meta-analysis of original data D_0 can now inform priors for use in EVSI calculations. However, the appropriate EVSI calculation depends crucially on the relationship between the target population in the decision problem and the M populations that have been studied so far. The main distinction to be made is between a decision that concerns a specific patient group, which may or may not have been studied before, and a decision relating to the entire population from which the M groups were drawn.

3.2. Collection of Further Information on a Specific Population

3.2.1. Target Population Studied Before

Information on population group K that was studied in D_0 could be used directly to inform a beta prior. However, as was implied above, a superior estimate of θ_K is obtained by considering this population as exchangeable with the other $M - 1$ populations.[48] This "shrinkage" toward an overall mean is especially beneficial if observations so far on the group have been extreme or based on sparse results. We therefore use ξ_K and τ_K as the prior mean and precision: $\xi = \text{logit}(\theta) \sim \text{Normal}(\xi_K, \tau_K)$.

3.2.2. The Target Population Has Not Been Studied Before

If the decision maker is prepared to consider the new population as "exchangeable" with the M populations studied so far, the relevant uncertainty can be expressed as the uncertainty relating to a new, $(M + 1)$th, patient population, drawn from the same overall distribution as the previous M patient populations. We can generate a predictive distribution by sampling from the population distribution.

$$\xi_{M+1} \sim \text{Normal}(\mu, \tau_\xi).$$

Exchangeability is a similar but weaker condition to be identically independently distributed. In this context, it means the each of the $M + 1$ patient populations is an equally likely draw from this common distribution.

3.2.3. Algorithms Based on Normal/Normal Conjugacy via Normal Approximation

The approach adopted here is to treat the prior, once it has been derived from the meta-analysis, as if it were based on a single previous study. Then the methods of section 2 are applicable, except that the prior information is not on the parameter θ itself but on a function of θ, $\xi = \log(\theta/(1-\theta))$. At this point, the discussion can be generalized to also cover random effects models for rates $\xi = \log\theta$, based on an equivalent normal theory hierarchical model as outlined above or a random effects model for some normally distributed continuous parameter. The appropriate procedure is then similar to what has been outlined for transformed continuous variables. The latter case is identical to the known variance situation in section 2. For rates and probabilities, we consider again EVSI of a study of size n and with F years of follow-up for studies on the rate parameter.

Algorithm for EVSI on a Parameter with a Log Odds or Log Prior

For each outer simulation *i:*

B1. (a) Draw a sample $\xi^{(i)}$ from its prior distribution Normal(μ_0, τ_0) (see above)

(b) Transform back to $\theta^{(i)}$:

logit of probability: $\theta^{(i)} = \exp(\xi^{(i)})/(1 + \exp(\xi^{(i)})$

log rate: $\theta^{(i)} = \exp(\xi^{(i)})$

B2. (a) Draw a sample from the distribution of the sufficient statistics:

Binomial numerator: $r_D^{(i)} \sim \text{Binomial}(\theta^{(i)}, n)$,

Poisson event count: $e_D^{(i)} \sim \text{Poisson}(\theta^{(i)} nF)$

(b) Find the sufficient statistics for sample, using the normal approximation.

logit probability: $\mu_D^{(i)} \approx \log(r_D^{(i)}/(n - r_D^{(i)}))$, $\tau_D^{(i)} \approx [1/r_D^{(i)} + 1/(n - r_D^{(i)})]^{-1}$

log rate: $\mu_D^{(i)} \approx \log(e_D^{(i)})$, $\tau_D^{(i)} \approx 1/e_D^{(i)}$

B3. (a) Update the prior with the new data to obtain parameters of the posterior distribution:

$$\xi \mid D^{(i)} \approx \text{Normal}((\mu_0\tau_0 + \mu_D^{(i)}\tau_D^{(i)})/(\tau_0 + \tau_D^{(i)}), \tau_0 + \tau_D^{(i)})$$

(b) For use in algorithms C2, C3, C4, transform this back to the original parameter scale:

logit of probability:

$\theta \mid D^{(i)} = \exp(\xi \mid D^{(i)})/(1 + \exp(\xi \mid D^{(i)}))$

log rate: $\theta \mid D^{(i)} = \exp(\xi \mid D^{(i)})$

(c) Alternatively, an approximate expression for the posterior expectation $E(\theta_I \mid D^{(i)})$ may be available in closed form for use in algorithm C1 (see Table 1).

3.3. Collecting Further Information on the Population Mean

A very different EVSI calculation is required if information is desired not on a single patient group but on the population mean μ itself. A situation in which this might be relevant would be, for example, a test sensitivity parameter that has been observed to vary unsystematically from center to center, perhaps as a result of variation in local expertise in its use. Changes in baseline treatment efficacy or disease progression parameters might also vary in this way. If this variation between centers was completely random, it might be reasonable to consider the M prior studies to be random draws from the national population. Then the overall mean μ would then be an estimate of the national average, and its precision τ_μ would appropriately represent the uncertainty in the mean. There are a number of important caveats to this analysis, which are examined later in this section.

To obtain new information on population average μ, it will not be useful to run a single study in 1 center. However large this study is, the addition of 1 more data point will do little to reduce τ_μ. Instead, a study is required in a number of centers, randomly chosen from the general target population. The design problem is to weigh the informational benefits and (low) costs of recruiting more individuals from each center with the higher costs and higher informational benefit of recruiting new centers. Of course, the number of centers must be reasonable to obtain a good estimate of the population mean, while the numbers per center must be high enough to give a reasonable estimate of the center-specific value. We will therefore examine how to calculate EVSI for any number m of centers and any number n of patients per center. This is illustrated here for an underlying probability parameter, but the results would follow through to any parameter for which a Gaussian hierarchical model can be constructed.

Algorithm for EVSI for the Mean of a Random Effects Distribution that Is Normal on the Logistic Scale

For each simulation *i:*

B1. (a) Draw m independent samples $\xi_k^{(i)}$, k = 1,2 . . . m, from Normal(μ, τ_ξ)
(b) Transform back to $\theta_k^{(i)} = \exp(\xi_K^{(i)})/(1 + \exp(\xi_k^{(i)}))$

B2. (a) Draw m binomial sample numerators, 1 for each k: $r_k^{(i)} \sim$ Beta($\theta_k^{(i)}$, n)
(b) Estimate the population average $\mu_D^{(i)}$ and its precision $\tau_{\mu D}^{(i)}$ from the m binomial samples using the closed-form moment estimator[29] as follows (temporarily dropping the simulation superscripts *(i)*)):

$$x_k = \text{logit}(r_k/n)$$

$$s^2 = \Sigma_k[x_k^2 - (\Sigma_k x_k)^2/m]/(m-1)$$

$$v_k = 1/r_k + 1/(n - r_k)$$

$$\sigma_\mu^2 = \max(0, s^2 - \Sigma_k v_k/m)$$

$$w_k = (v_k + \sigma_\mu^2)^{-1}$$

$$\mu D^{(i)} = \Sigma w_k x_k/\Sigma w_k,\ \tau_{\mu D}^{(i)} = \Sigma w_k$$

B3. (a) Update the prior distribution with the new sample data to obtain parameters of the post erior distribution:
$\xi \mid D^{(i)} \sim$ Normal($(\mu\tau_\mu + \mu_D^{(i)}\tau_{\mu D}^{(i)})/(\tau_\mu + \tau_{\mu D}^{(i)})$, $\tau_\mu + \tau_{\mu D}^{(i)}$)
(b) and (c) Proceed as for the algorithm for parameters with a log odds prior

A weakness in the approach is the use of the moment estimator, which can produce negative variances. This is one of several approximations that are taken up in discussion.

A word of caution is required regarding the interpretation of the random effects distribution for these purposes. First, there is no guarantee that the previous M studies have sampled from the overall patient population in a representative way. For example, if patient groups differed in the baseline probability of disease progression, then the choice of population of interest in the M previous RCTs might be far from random. If, however, interest was focused on a wider population that perhaps included some or all the previous M, then it might be possible to reconstruct the required target population as a suitably weighted μ and its precision τ_μ using the ξ_i and τ_i from the prior meta-analysis with population sizes as weights. Uncertainty regarding the weights could also be built in using a Dirichlet distribution.

A 2nd difficulty with the approach we have suggested is that there is a danger of confusing parameter uncertainty with parameter variation. If the reason for parameter heterogeneity is known, for example, an association with age or disease severity, then the decision problem should possibly be focusing on the age or severity threshold at which one treatment becomes cost-effective relative to another. Uncertainty regarding the threshold could be reduced by further research, but this leads to questions concerning the uncertainty in the mathematical relationship between θ and the covariate, which takes us beyond the scope of this article.

4. UNCERTAINTY IN MEASURES OF RELATIVE EFFICACY

4.1. Relative Measures Involving Probabilities: Risk Difference, Risk Ratio, and Odds Ratio

In comparative cost-effectiveness analyses involving clinical interventions, the decision model includes 2 parameters, θ_1 and θ_2, representing, say, the probabilities of recovery on treatments 1 and 2. In a "textbook" presentation of such a problem, one would assume Beta priors for each parameter. However, almost invariably, both the epidemiological literature and expert opinion on θ_1 and θ_2 will be expressed in terms of some relative efficacy measure—a risk difference, risk ratio, or odds ratio—and these same measures will be employed in the design and analyses of planned trials. (See Deeks[49] for discussion of the merits of alternative relative effect measures.) The decision question is to choose an optimal allocation $\{n_1, n_2\}$ for a randomized trial. It is assumed, in addition, following common practice in decision literature, that the baseline recovery rate, θ_1, has already been determined, either from other trials or from a cohort study. Often, θ_1 is taken to be a constant. However, to preserve generality, we assume it is a random variable with a beta or gamma prior, and furthermore, the prior estimate of θ_1 is independent of the prior for the relative effect measure. The proposed RCT will not update this parameter, only the relative effect measure. Alternatives are examined in section 4.3. The general form of the algorithm in this situation is as follows, with the appropriate formula for risk difference (RD), log odds ratio (LOR), log risk ratio (LRR), and log hazard ratio (LHR).

Table 2 Expectations of Random Variable $E(\theta_2)$ Given the Distributions of θ_1 and Relative Effect Measures $G(\theta_1, \theta_2) \sim \text{Normal}(\mu, \tau)$, Assuming θ_1 and $G(\theta_1, \theta_2)$ Are Independent

Relative Effect Measure	$G(\theta_1, \theta_2) \sim N(\mu, \tau)$	θ_1	$E(\theta_2)$
Risk difference	$\theta_2 - \theta_1$	Beta(a, b)	$\mu + a/(a + b)$
Log odds ratio*	$\log(\theta_2(1 - \theta_1)/(\theta_1(1 - \theta_2)))$	Beta(a, b)	$h + (1/\tau + 1/a + 1/b)h(1 - h)(1 - 2h)/2$, where $h = e^w/(1 + e^w)$, and $w = \mu + \log((a - 0.5)/(b - 0.5))$
Log relative risk*	$\log(\theta_2/\theta_1)$	Beta(a, b)	$a \exp(\mu + 1/(2\tau))/(a + b)$
Log hazard ratio*	$\log(\theta_2/\theta_1)$	Gamma(a, b)	$a \exp(\mu + 1/(2\tau))/(a + b)$

*Second-order Taylor series expansion.[68]

Algorithm for EVSI for Relative Effect Measures

For each simulation *i:*

B1. (a) Draw a sample from the prior distribution $\xi^{(i)} = G(\theta_1, \theta_2) \sim \text{Normal}(\mu_0, \tau_0)$ of the relative effect measure, and draw a sample baseline parameter $\theta_1^{(i)}$ from its prior distribution.

RD, LOR, LRR: $\theta_1^{(i)} \sim \text{Beta}(a, b)$

LRH: $\theta_1^{(i)} \sim \text{Gamma}(n_0, Y_0)$

(b) Transform back to obtain an implied prior for $\theta_2^{(i)}$:

RD: $\theta_2^{(i)} = \xi^{(i)} + \theta_1^{(i)}$

LOR: $\theta_2^{(i)} = \exp(\xi^{(i)})/(1 - \theta_1^{(i)} + \theta_1^{(i)} \exp(\xi^{(i)}))$

LRR: $\theta_2^{(i)} = \theta_1^{(i)} \exp(\xi^{(i)})$

LHR: $\theta_2^{(i)} = \theta_1^{(i)} \exp(\xi^{(i)})$

B2. (a) Draw a sample sufficient statistic $D^{(i)}$, a binomial numerator or Poisson event count as appropriate, for each arm t,

RD, LOR, LRR: $r_t^{(i)} \sim \text{Binomial}(\theta_t^{(i)}, n_t)$

LHR: $e_t^{(i)} \sim \text{Poisson}(\theta_t^{(i)} n_t F)$

(b) Convert the sufficient statistics to a mean and variance using the normal approximation:

RD: $\mu_D^{(i)} = r_2^{(i)}/n_2 - r_1^{(i)}/n_1$, $\tau_D^{(i)} \approx [r_2^{(i)}(n_2 - r_2^{(i)})/n_2 + r_1^{(i)}(n_1 - r_1^{(i)})/n_1]^{-1}$

LOR: $\mu_D^{(i)} = \log[(r_2^{(i)}(n_1 - r_1^{(i)})/(r_1^{(i)}(n_2 - r_2^{(i)}))]$,

$\tau_D \approx [1/r_1^{(i)} + 1/(n_1 - r_1^{(i)}) + 1/r_2^{(i)} + 1/(n_2 - r_2^{(i)})]^{-1}$

LRR: $\mu_D^{(i)} = \log[r_2^{(i)} n_1/(r_1^{(i)} n_2)]$,

$\tau_D^{(i)} \approx [(n_1 - r_1^{(i)})/(n_1 r_1^{(i)}) + (n_2 - r_2^{(i)})/(n_2 r_2^{(i)})]^{-1}$

LHR: $\mu_D^{(i)} = \log[e_2^{(i)} n_1/(e_1^{(i)} n_2)]$, $\tau_D^{(i)} \approx [1/e_1^{(i)} + 1/e_2^{(i)}]^{-1}$

B3. (a) Update the prior with the new sample to obtain parameters of the posterior distribution:

$\xi \mid D^{(i)} \sim \text{Normal}((\mu_0\tau_0 + \mu_D^{(i)}\tau_D^{(i)})/(\tau_0 + \tau_D^{(i)}), \tau_0 + \tau_D^{(i)})$

(b) For use in algorithms C2, C3, C4, back transform to obtain $\theta_2 \mid D^{(i)}$ using the relationships set out in step B1(b) and $\theta_1^{(i)}$

(c) If conditions are met for algorithm C1, use the approximations for the posterior means $E(\theta_2 \mid {}_D{}^{(i)})$ in Table 2.

4.2. Heterogeneity of Prior Information in Measures of Relative Efficacy

The relative merits of fixed or random effects meta-analysis in the context of relative efficacy measures have been well rehearsed.[30,31,50] The comments in section 3 on heterogeneity of proportions or other absolute quantities apply equally to relative effect measures. In other words, the decision maker must determine whether the target population has been studied before and if not, whether the new group is exchangeable with groups that have been studied or whether interest centers on the average of a set of patient groups. The methods outlined previously can then be applied using the formulae above for updating relative effect measures and normal approximations for the variances of the relevant relative effect measures.

However, the danger of confusing uncertainty with variation appears to be greater in a relative effect measure than in a single parameter, and the nature of variation and its implications for the target population must be considered carefully. There may certainly be scope for considering a national average LOR, when it is known that relative treatment efficacy varies unsystematically between centers, perhaps because of the

way one of the treatments is administered. However, this seems intuitively much less likely than unsystematic variation in baseline parameters. Metaregression[51–53] is the statistical technique for determining the relationship between relative treatment efficacy and a covariate such as age that may explain heterogeneity. Once again, there is every possibility that further information is required on the relation between relative efficacy and the covariate.

4.3. Updating Both Baseline and Relative Effect Measures

Whether or not the prior for baseline efficacy, θ_1 on treatment 1, is based on cohort or RCT studies, it may make sense to use results of a new study to update the baseline parameter as well as the relative effect measure. The comparator treatment arm of the trial will, after all, represent the efficacy of placebo or standard treatment under the relevant, contemporary circumstances. Technically, the algorithm of the previous section can be easily modified to achieve this: All that is required is that the simulated data in step B2 is used to update not only the relative effect measure but also the baseline efficacy via the beta binomial in section 2. Then, in step B3, the posterior distribution is used in steps B3(a, b) or its parameters in steps B3 (c).

However, this procedure has an interpretation that may not be the intended one: It assumes that there is a single, constant efficacy under standard care that does not vary between studies and that can therefore be updated by additional information. This would depart from meta-analysis practice, in which the study effect is usually removed by using the relative effect measures as primary data or may at best be modeled as a random effect. If this line of reasoning is accepted, then the EVSI calculations for studies that update the baseline study effect as outlined above will be valid only if the decision problem defines a specific patient group and some specific circumstances for which it can be reasonably held that the baseline really is fixed. If a fixed effects model cannot be assumed, then the baseline should be considered to be a random effect, and if further information is needed, then this should be further information on its mean (see section 3.3).

5. DISCUSSION

EVSI calculations can be carried out for 1 or more parameters, in a wide range of decision models, in the presence of multiple sources of uncertainty, and with priors structured in the form typically seen in epidemiological literature. This article has shown how these methods can be applied when prior information informs us about transformations of the parameters that appear in the model and, in particular, the canonical situation with priors on relative efficacy measures. In addition, the prior information available on parameters may exhibit additional sources of uncertainty or variation (heterogeneity), over and above sampling error. This has led us to propose different EVSI calculations relating to the collection of further information on a previously studied group, on a new "exchangeable" group, or on the average of a set of heterogeneous population groups.

5.1. The Scope of Proposed Methods

Expected value of information theory is concerned with the accurate quantification of uncertainty. The overriding feature of these methods is therefore that they can be applied only to fully probabilistic models. Although many deterministic models can be converted into probabilistic form, those that rely on Monte Carlo simulation to be evaluated for a fixed set of parameter inputs require a further outer simulation even to determine the decision maximizing expected net benefits under parameter uncertainty. In many cases, an additional level of Monte Carlo simulation may prove computationally nonfeasible. This is not, of course, a failure of EVSI theory or a limitation of the algorithms proposed here but simply a reflection of the computational cost of microsimulation and discrete event simulation, which encourages reliance on deterministic forms of sensitivity analysis. However, Monte Carlo simulation is not the only way of evaluating probabilistic models: Gaussian process models can be used to produce more accurate estimates with far fewer simulations.[54–57]

Given a computationally manageable probabilistic model, the EVSI algorithms we have proposed are flexible, accommodating parameters in complex relations to the data to be collected and nonlinear net benefit functions. In effect, the preposterior uncertainty is propagated through the decision model in the same way as prior uncertainty in the evaluation of prior net benefit. Where net benefit is linear in all parameters, Taylor series approximations are often available, allowing evaluation via a single Monte Carlo cycle even when information is obtained on nonlinear functions of parameters.

There are, however, significant limitations. One important class of restrictions are those regarding prior parameter correlation. If there are correlations between the parameters θ_I of interest that are to be updated and the remaining set θ_I^c, then new data on θ_I will provide information about θ_I^c, and the calculations of posterior net benefit, including the mean posterior net benefit, will be incorrect. (We exclude from this the specific case of prior correlation that can be represented as a multivariate normal distribution, as this can be readily updated with a multivariate normal likelihood.) It could be argued that the restriction is not significant as the vast majority of decision models are populated with independent parameters derived from unrelated literature sources. However, this way of populating decision models owes a lot to the philosophy behind deterministic modeling. The uncorrelated priors assumption may become less relevant, as the whole thrust of probabilistic, Bayesian, decision analysis—and above all, EVI theory—requires that uncertainties are accurately represented and therefore based on all available data, rather than simply "enough" data to populate the model. Incorporation of all the available data will often necessitate use of computational techniques for combining information on parameters with information on model outputs and other complex functions of several parameters. Such methods inevitably induce parameter correlation. These include the confidence profile method of David M. Eddy and his colleagues[58] and the Bayesian Monte Carlo methods in the risk analysis literature.[59–61] The ready availability of flexible Bayesian Markov chain Monte Carlo (MCMC) software to carry out these and still more complex multiparameter evidence syntheses[62] is likely to accelerate a trend away from simple uncorrelated prior parameter structures.

Another related limitation is that priors must be represented in a fully parametric form. Although a joint posterior distribution from a Bayesian MCMC analysis of current data would conceptually represent the ideal prior for EVSI calculations, this does not constitute a set of priors that are amenable to updating with conjugate likelihoods. There are, in fact, few examples of EVSI calculations that have not relied on conjugacy to achieve computational feasibility, but one exception is Bayesian Monte Carlo.[14] This is one of a class of several nonparametric, noniterative Monte Carlo techniques[48] for updating a prior with new data. In the EVSI context, these methods require nested Monte Carlo simulation but could be applied with correlated priors, or when the new data induces parameter correlation, and in nonlinear models. Noniterative MC methods have the general difficulty that they can be inefficient if the prior parameter space and the likelihood parameter space are disparate, and it is therefore too early to determine whether they can be used routinely in EVSI as application-specific modifications may be required.[61] The increasingly popular MCMC methods avoid these problems by converging quite rapidly toward the posterior parameter space, but the fact that the first few thousand "burn in" simulations have to be discarded presents a barrier to their routine use in the inner integration step of EVSI calculations. There are, however, a wide range of computational methods for Bayesian updating,[48,63] which, although somewhat sidelined by MCMC for general Bayesian data analysis, could still be highly relevant in computation of EVSI in medical applications. This is an area for future research.

5.2. Approximate Nature of Proposed Solutions

Our main objective in the article has been to clarify the nature of EVSI calculations in an epidemiology context for a broad range of parameters and information structures. Simplifying assumptions have therefore been made to preserve generality and facilitate exposition. Having said this, the proposed algorithms are intended for use in real applications, so it is worth considering the potential for improvement.

Apart from the algorithms proposed in section 2, the methods used here must be regarded as only approximate. Normal approximations are used in many places to create the normal-normal conjugacy conditions that allow binomial and Poisson distributed data to be combined with priors that are commonly expressed in terms of normal distributions, and it is not known what effect this might have on accuracy. Although there are "exact" methods for updating priors on LOR,[64] and even risk ratios and risk differences with binomial data,[65] these generally require MCMC.

In basing priors on the output from the meta-analysis of existing data set D_0, a series of additional approximations have been made. First, it was assumed that the population variance σ_ξ^2 was known. It is well known that this assumption can create problems: Negative variance estimates can occur, and there is a tendency to exaggerate the uncertainty in extreme ξ_i. The difficulty does not lie so much in representing the possibility that the population variance is unknown as much as in reflecting the complex nature of the joint uncertainty of σ_ξ^2, μ, and the study specific true means ξ_i, while still devising a simple algorithm for updating with new data D.

Another approximation, in dealing with heterogenous priors, was that all the prior information was summarized in terms of parameters of normal distributions for μ and the ξ_i. New data D was used to update these distributions, but at this point, the original meta-analysis structure had been discarded. An alternative would be to analyze the new sample $D^{(i)}$ on each simulation i along with previous data D_0 as the $(M + 1)$th group. This seems more logically consistent, as the posterior analysis post $D^{(i)}$ is exactly the same as the earlier posterior analysis post D_0. Compared to this "re-meta-analysis," it is likely that the methods we have proposed will overestimate the value of the new study on the new $(M + 1)$th patient population, especially for large M. This is because if the preposterior analysis were based on the full $M + 1$ group data set, the mean for group $M + 1$ would be drawn in toward the population mean. The methods proposed by Abrams and Sanso,[66] based effectively on the same 2-stage Gaussian model for meta-analysis assumed in section 3, provide accurate approximations for posterior means of μ and ξ_i that can be computed in closed form. Suitably adapted, these could have an important role in some EVSI calculations.

It should not be forgotten, of course, that Monte Carlo methods for integration are themselves only approximations, and the appropriate number of simulations from the prior and nested "inner" Monte Carlo simulations for algorithms C2, C3, and C4 is itself a nontrivial matter. Bearing in mind that the EVSI calculation is set within an optimization problem for sample size, the attractions of the C1 algorithm that relies only on prior and posterior means on each outer simulation are all the stronger. We have recommended the use of Taylor series approximations when EVSI requires updating nonlinear functions of parameters in otherwise linear models. But it is possible that this approach could sometimes also be valuable when net benefits are not linear in parameters.

The accuracy of the outer simulation depends on the number of iterations on which the strategy option with the maximum postsample net benefit differs from the current optimal strategy. This can be assessed from the mean and variance of the prior incremental net benefits. Furthermore, when plugging parameter means or subtracting the expected value of the current decision, it is also essential to ensure these were generated by the same random number sequence. This will reduce the error due to simulation. Whenever the net benefit is not linear in all its parameters, nested Monte Carlo calculations are required, but the relative effect of increasing the number of inner or outer simulations is not known. The Gaussian process approach to computationally expensive models could again be useful.[54]

5.3. Optimal Expected Net Benefit of Sampling as the Context for EVSI

It is worth reminding ourselves that EVSI calculations are set within an optimization problem for expected net benefit of sampling. In section 1.6, for example, the case of P parameters passing the EVPPI "hurdle" was mentioned, as was the requirement to optimize over the vector of samples sizes $n = \{n_1, n_2, \ldots n_P\}$. Recent work in the hydrology literature has compared 2 general optimization methods for such problems, the branch and bound and the genetic algorithm.[13] However, this formulation of the problem, in which each unknown parameter is informed by an item of data, will not always be applicable in a medical context and does not seem to entirely cover what one might mean by an "optimal research portfolio." As we have seen, data on a relative effect measure—a function of parameters—may be required, as well as or instead of a baseline. At the same time, a single study might provide information on the proportion of patients able to benefit from a treatment, on a relative probability of avoiding a defined illness condition, on baseline treatment efficacy, on the time spent in hospital by those who become ill, on average utilities in certain states, and on certain cost components.[67] The precise characterization of the real research space and mathematical solutions to the optimization problem are issues for further research.

5.4. Conclusion

This article has attempted to set out a technical approach to EVSI calculations that fits within the Monte Carlo computation framework generally used for probabilistic decision analysis. However, to develop computational methods relevant for the form in which epidemiological data are analyzed and presented in the literature, it has been necessary to explore many nontechnical issues relating to the interpretation of epidemiological data. In particular, whether the uncertainty that further data are supposed to reduce is uncertainty in the absolute parameter values or in a relative effect measure. In addition, EVI analysis requires investigators to be completely specific about all parameter uncertainty relevant to the decision and reveals searching questions about the origin of uncertainty and variation in epidemiological data, especially the relationship be-

tween the target population for decision and the populations previously studied. These are not new questions, but the EVI analysis exposes the fact that they must be answered if coherent research planning and design are to be undertaken.

We have shown how EVSI calculations can be carried out via a single or double Monte Carlo simulation for a very wide range of decision situations, given an explicit model for all sources of uncertainty and variation. Further research is required to extend the availability of these powerful decision theoretic calculations to a still wider class of models, to characterize the concept of an optimal research portfolio, and to suggest solutions to the optimization problems that this represents.

APPENDIX
Illustrative Calculations for Expected Value of Sample Information (EVSI) in a Simple Decision Model

This appendix provides practical examples of several of the algorithms mentioned in the main text. We begin by presenting a fictitious decision model, then derive the optimal current decision and its value. A computing strategy is outlined before we set out EVSI calculations for different parameters. In the appendix, we simply outline the main features of the calculations along with some illustrative results. SAS code for all the examples is available at http://www.hsrc.ac.uk/Current_research/research_programmes/mpes.htm.

The Model

A new treatment has been proposed for an illness in which a critical event occurs with probability p_C on standard care. The new treatment is thought to reduce the risk of the critical event substantially but it is accompanied by side effects, with probability p_{SE}. There is prior information on the efficacy of the new treatment in the form of a log odds ratio (LOR) and also on the probability of side effects, but both are limited. The critical event leads to a decrement in quality of life, Q_E, but this too has been assessed imprecisely. The decision model is shown in Figure A1. Its parameters and their prior expectations and distributions appear in Table A1.

The A Priori Decision and Its Expected Value

Using the model in Figure A1, we can readily set out expressions for the net benefit $B(t, \theta)$, with t taking the values C for the standard care and T for the new treatment.

Net benefit equations are as follows:

$$nb_C = p_C * (-C_E + W * L * (1 + Q_E)/2) + (1 - p_C) * W * L$$

$$nb_T = -C_T + p_{SE} * p_T * (-C_{SE} - C_E + W * (-Q_{SE} + L * (1 + Q_E)/2)) + p_{SE} * (1 - p_T) * (-C_{SE} + W * (L - Q_{SE})) + (1 - p_{SE}) * p_T * (-C_E + W * L * (1 + Q_E)/2) + (1 - p_{SE}) * (1 - p_T) * W * L.$$

Below, we will write it as a known function of 5 arguments, $B(t, p_{SE}, Q_E, p_C, p_T)$.

Table A1 Model Parameters, Prior Distributions, and Their Means

Description	Parameter	Mean	Distribution
Mean remaining lifetime	L	30	Constant
QALY after critical event, per year	Q_E	0.6405	logit(Q_E) ~ N(0.6,6)
QALY decrement due to side effects	Q_{SE}	1	Constant
Cost of critical event	C	$200,000	Constant
Cost of treatment	C_T	$2000	Constant
Cost of treatment side effects	C_{SE}	$100,000	Constant
Monetary value of 1 QALY	W	$75,000	Constant
Probability of critical event, no treatment	p_C	0.15	Beta(15,85)
Probability of treatment side effects	p_{SE}	0.25	Beta(3,9)
Odds ratio, $p_T(1 - p_C)/(p_C(1 - p_T))$	OR		log(OR) ~ N(–1.5,3)
Probability of critical event on treatment	p_T	0.0440	

Note: QALY = quality-adjusted life year. The parameters of normal distributions are given as mean and precision (= 1/variance). Prior means of parameters Q_E and p_T are based on 100,000 simulations.

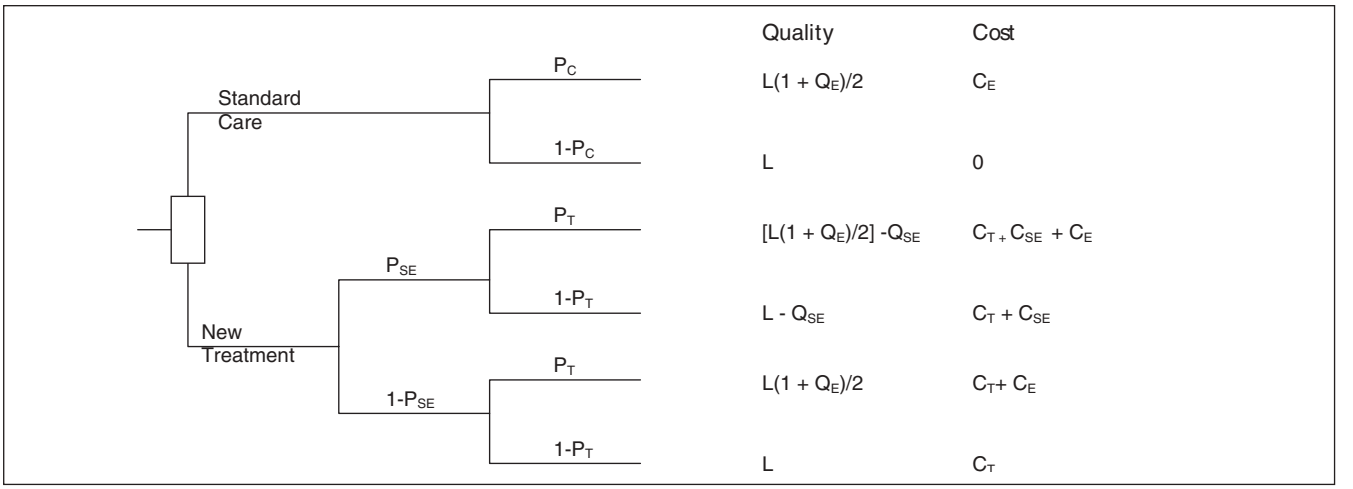

Figure A1 Decision model. See Table A1 for prior expectations and distributions.

The equations are multilinear in all their parameters so that the expected net benefits can be found by plugging in their mean values. Care must be exercised here as we do not have the prior means of Q_E and p_T immediately available but must obtain them indirectly, either through simulation or using the Taylor series formulae in Tables 1 and 2. As explained below, we chose to calculate the expected net benefit by simulation, deriving values of Q_E and p_T on each Monte Carlo cycle via the following relationships:

$$Q_E = \exp(\text{Logit}(Q_E)/(1 + \exp(\text{Logit}(Q_E))$$

and

$$\text{logit}(p_T) = \log(p_C/(1 - p_C)) + LOR$$

$$p_T = \exp(\text{Logit}(p_T)/(1 + \exp(\text{Logit}(p_T)).$$

Based on 100,000 simulations, we find expected net benefits are $2,159,300 for standard care and $2,164,700 the new treatment, to the nearest $100. The prior decision is therefore strategy *T*, new treatment, and the expected value of this decision is $2,164,700.

Computing Strategy

Because the algorithms we propose are all based on Monte Carlo simulation, results are subject to simulation error. When drawing samples from a prior distribution beta(15,85), there is no guarantee that the mean of even 100,000 simulations will be exactly the 0.15 expected. In addition, the final step of all EVI calculations involves subtracting the expected value of the current optimal decision from the expected value of a decision based on perfect or sample information. If these 2 elements are calculated from different random number sequences, further inaccuracy will occur.

The following computing strategy has been adopted to mitigate these problems:

1. For each prior distribution, the same random number sequence is used for all calculations.
2. When the prior expectation of a parameter is required, the value used is the mean given by that random Monte Carlo sequence, which will be called θ_0.
3. We rewrite EVSI expression (4) from section 1.5,

$$E_D \max_t E_{\theta_I^c,(\theta_I \mid D)} B(t, \theta_I, \theta_I^c) - \max_t E_{\theta} B(t, \theta_I, \theta_I^c),$$

as

$$E_D\left(\max_t E_{\theta_I^c,(\theta_I \mid D)} B(t, \theta_I, \theta_I^c) - E_{\theta_I^c,(\theta_I \mid D)} B(t^*, \theta_I, \theta_I^c)\right),$$

where t^* indicates the a priori optimal strategy, in this case *T*. This means that on each draw from the prior, we compute the expected value of both decisions based on the predicted posterior net benefits and record either a zero when *T* is the posterior decision or the difference between the expected posterior benefits when *C* is best. This partially controls for the effect of the random Monte Carlo sequence used to draw sample statistics from their predictive distributions (step B2) and also for the effects of any normal approximations that are used to derive sample statistics in a form that is conjugate with the priors.

Expected Value of Perfect Information (EVPI)

Using the formulation above, EVPI is the expected difference between the value of the optimal decision based on perfect information and the value of the prior decision. We take the maximum of $B(C, \theta^{(i)})$ and $B(T, \theta^{(i)})$ on each Monte Carlo cycle, subtract $B(T, \theta^{(i)})$, and average over all simulations. The EVPI per patient turns out to be $10,140. We can estimate the probability that the optimal prior decision will turn out to be the wrong decision by counting the number of simulations in which $B(C, \theta^{(i)}) > B(T, \theta^{(i)})$. This probability is 0.43.

EVSI for Probability of Side Effects, p_{SE}

What would it be worth to conduct an observational study on $n = 60$ patients who are on the new treatment? Following the algorithm in section 2.1, we draw 100,000 values from the prior distribution $p_{SE} \sim$ beta(3,9) (step B1), then for each $p_{SE}^{(i)}$, we draw a binomial sample $r_D^{(i)}$ from Binomial($p_{SE}^{(i)}$, $n = 100$) (step B2). For each simulated study numerator $r_D^{(i)}$, the posterior mean $p_{SE}^{(i)}$ is given by

$$E\,[p_{SE} \mid D^{(i)}] = (a + r_D^{(i)})/(n_0 + n),$$

where $a = 3$, $n_0 = 12$, and $n = 60$. The above posterior mean value of p_{SE} on each Monte Carlo cycle i is then used in the formula for net benefit (step B3), with all the other parameters taking their prior mean value (algorithm C1) as net benefit is linear in all parameters. Specifically, we find the value of

$$E\,[B(\,t, \theta \mid D^{(i)})] = B\,(t, E\,[p_{SE} \mid D^{(i)}], E[Q_E], E\,[p_C], E\,[p_T])$$

$$= B(t, E\,[p_{SE} \mid D^{(i)}], 0.6402, 0.1499, 0.0440).$$

For both $t = C$ and $t = T$ on each cycle i, find the maximum and subtract from it the value of the current decision, that is, where $t = T$. The average of this series of zeros and differences is the EVSI, which is \$5,550 for a study of size $n = 60$. (The constants in the equation are the values θ_0 of the prior expectations as computed by simulation.)

Because side effects do not occur under standard care, the term p_{SE} does not appear in the equation for net benefit under C, and this expression is a constant over each simulation. Values of $B(T, E\,[p_{SE} \mid D^{(i)}\,], 0.6402, 0.1499, 0.0440)$ that are less than \$2,159,300, the expected benefit under C, will therefore lead to a change in decision. Figure A2 shows the distribution of $E[B(t, \theta \mid D^{(i)}\,)] - 2{,}159{,}300$ across the simulations. This is the so-called preposterior distribution of incremental net benefit (see section 2.1). An additional study of $n = 60$ patients changes our prior decision in 34% of simulations, the area to the left of zero. The discrete nature of the distribution in Figure A2 is due to the fact that the only variable in $E[p_{SE} \mid D^{(i)}]$ is an integer $r_D^{(i)}$.

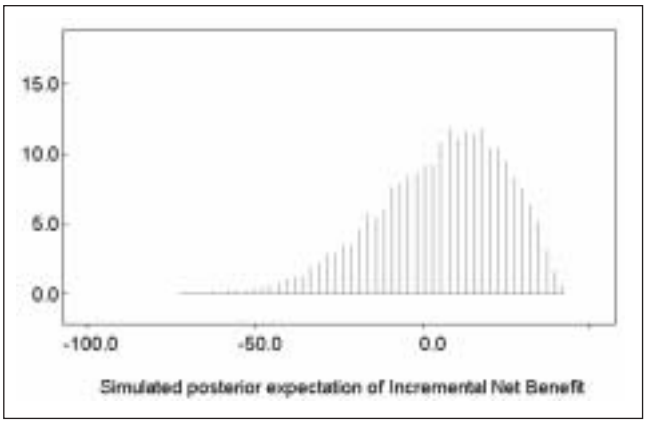

Figure A2 Histogram of the simulated posterior expectation of incremental net benefit given further information on the parameter p_{SE} *based on a sample of* $n = 60$*, also known as the preposterior distribution. For samples on the left of zero (34% of occasions), the expected net benefit after the new sample is greater under standard care, and a decision maker would switch treatment.*

It is instructive to observe the effect of different sample sizes. Table A2 shows how EVSI increases with sample size. Because of the relatively low precision of the prior mean, even modest study sizes contribute quite substantially to the decision. The table also gives the probability that the new sample will change the initial decision, that is, $Pr\{B(C, E[\theta \mid D^{(i)}])\ B(T, E[\theta \mid D^{(i)}])\}$. This appears at first sight to behave inconsistently, rising from 0.25 with a sample size of 1 and

Table A2 Expected Value of Sample Information (EVSI) Calculations for p_{SE} with a Range of Sample Sizes, and Probability That the New Study Will Change the Current Decision and Preposterior Means for p_{SE} and for Net Benefit under New Treatment

n	EVSI (\$)	Probability of Change in Decision	Preposterior s (p_{SE})	Preposterior s ($B(T, \theta)$
1	1190	0.25	0.03	5.8
5	2750	0.37	0.07	11.4
10	3630	0.27	0.08	14.2
20	4550	0.39	0.09	16.6
40	5250	0.36	0.11	18.4
60	5550	0.34	0.11	19.1
100	5820	0.36	0.11	19.8
200	6010	0.36	0.12	20.3
500	6150	0.36	0.12	20.7
1000	6190	0.37	0.12	20.8
10,000	6240	0.37	0.12	20.9
10,000,000	6240	0.37	0.12	20.9

Note: These are the distributions of the predicted posterior means over the 100,000 Monte Carlo draws.

then rising to 0.36 with $n = 5$, but falling again with $n = 10$. This is again due to the discrete nature of the posterior distribution of p_{SE}. With $n = 1$, for example, based on the prior mean of 0.25 for p_{SE}, the new study will produce $r_D = 1$ in 0.25 simulations. On these occasions, the posterior mean for p_{SE} must be $(3 + 1)/(12 + 1) = 0.308$, which is high enough to switch the decision to standard care. Table A2 also shows how the preposterior variance of p_{SE} and of net benefit under new treatment, $E[B(T, \theta \mid D^{(i)})]$, increase with increasing sample size.

EVSI for Quality of Life after Critical Event, Q_E

The prior distribution for logit(Q_E) is normal, with a mean of 0.6. Assume this was based on a study of $n_0 = 12$ patients with side effects. A decision maker now wants to assess the EVSI attaching to a new study on $n = 100$ new patients. The precision (1/variance) of the prior distribution is $n_0\tau_0 = 6$ (see Table A1), assumed to be known, therefore the population precision $\tau_0 = 0.5$.

We again take 100,000 simulations (A1), and on each simulation, we draw a sample $lQ_E^{(i)} = \text{logit}(Q_E)^{(i)}$ from its prior distribution $N(0.6,6)$ (step B1), and then a new sample mean $\mu_D^{(i)}$ from $N(lQ_E^{(i)}, 0.5n)$ (step B2). The precision of the sample mean, $0.5n$, is based on the population precision, 0.5, and the size of the proposed new study, n. Next (step B3), the prior is updated with the new sample to give a posterior expectation:

$$ElQ_E^{(i)} = E\,[\text{logit}(Q_E) \mid D^{(i)}] = (0.6 \times 6 + \mu_D^{(i)} \times 0.5n)/(6 + 0.5n).$$

Although the net benefits are linear in all parameters, we cannot immediately plug in a posterior mean for Q_E because at this point, we have only a posterior mean for logit(Q_E). However, on each simulation, we have a value for both the mean and precision $(6 + 0.5 \times 30)$ for the posterior distribution $\text{logit}(Q_E) \mid D^{(i)}$, and we can use the Taylor series method (see Table 1) to obtain an approximate posterior mean for Q_E on each cycle, as follows:

$$h^{(i)} = \exp(ElQ_E^{(i)})/(1 + \exp(ElQ_E^{(i)}))$$

$$E[Q_E \mid D^{(i)}] = h^{(i)} + h^{(i)}(1 - h^{(i)})\,(1 - 2h^{(i)})/\,2(6 + 0.5n))$$

and plug this into the net benefit equations, with remaining parameters taking the value of their prior means

$$B(t, E[p_{SE}], E[Q_E \mid D^{(i)}], E[p_C], E[p_T])$$

$$= B(t, 0.2498, E[Q_E \mid D^{(i)}], 0.1499, 0.0440).$$

Averaging over all simulations, we find that a new study of Q_E with $n = 100$ patients has an EVSI of $1880.

EVPI on the LOR

What would be the expected value of a decision based on perfect information about the LOR? Methods for the expected value of partial perfect information (EVPPI) in general have been given by Felli and Hazen[9]; however, as LOR is a nonlinear function of parameters, these methods need to be slightly modified. We provide this illustration here as it allows us to compare 1-stage EVSI computation with 2-stage EVSI, illustrates the use of the Taylor series formulae in Table 2, and serves as an introduction to the EVSI calculations for LOR.

To compute EVPPI for LOR, we must draw repeated samples $LOR^{(i)}$ from the prior for LOR, and for each sample, we must identify the maximum net benefit. First, imagine that we set out to evaluate expected net benefit on cycle i by a 2nd stage, inner Monte Carlo simulation. On this inner cycle, we would undertake the following calculations on each inner cycle k nested within i:

1. Draw a sample $p_{SE}^{(i,k)}$ from the prior distribution beta(3,9).
2. Draw a sample $\text{logit}(Q_E^{(i,k)})$ from its prior distribution $N(0.6,6)$.
3. Calculate $Q_E^{(i,k)} = \exp\,(\text{logit}(Q_E^{(i,k)}))/(1 + \exp(\text{logit}(Q_E^{(i,k)})))$.
4. Draw a sample $p_C^{(i,k)}$ from its prior beta(15,85).
5. Calculate $\text{logit}(p_T^{(i,k)}) = \text{logit}\,(p_C^{(i,k)}) + LOR^{(i)}$.
6. Calculate $p_T^{(i,k)} = \exp\,(\text{logit}(p_T^{(i,k)}))/(1 + \exp(\text{logit}(p_T^{(i,k)})))$.
7. Calculate $B(C, \theta^{(i,k)} \mid LOR^{(i)})$, and $B(T, \theta^{(i,k)} \mid LOR^{(i)})$ by plugging in the above values into the net benefit equations. Note that step 5 draws a sample from the conditional distribution of logit(p_T) conditioning on a known value of $LOR^{(i)}$.

After running through all the cycles k, we compute the average net benefits and calculate

$$\text{Max}_t\,(E[B(C, \theta \mid LOR^{(i)})], E[B(T, \theta \mid LOR^{(i)})] - E[B(t^*, \theta \mid LOR^{(i)})].$$

This process is repeated over the outer Monte Carlo simulations i, and the average of this expression is the EVPPI for LOR alone. With 10,000 inner simulations and 5000 outer cycles, we obtained EVPPI = $3890.

Now compare this algorithm to the 1-stage version made feasible because the model is multilinear in its parameters. We can take advantage of this by plugging in the appropriate expectations into the net benefit equations, rather than computing them by simulation. In particular, we can plug in the prior expectations of Q_E, p_{SE}, and p_C and the conditional expectation of $p_T \mid LOR^{(i)}$. To obtain the latter, we use the mean and variance of the conditional distribution of $\text{logit}(p_T) \mid LOR^{(i)}$. The required conditional expectation of p_T on each cycle i is computed as follows (see Table 2):

$$Elp_T^{(i)} = \log(14.5/84.5) + LOR^{(i)}$$

$$h^{(i)} = \exp(Elp_T^{(i)})/(1 + \exp(Elp_T^{(i)}))$$

$$Ep_T^{(i)} = h^{(i)} + h^{(i)}(1 - h^{(i)})(1 - 2h^{(i)})\text{Var}(lp_T^{(i)})/2,$$

where, as can be seen from the 2-stage algorithm outlined earlier, $\text{Var}(lp_T^{(i)})$ is given by

$$\text{Var}(lp_T^{(i)} = \text{Var}(lpc) + \text{Var}(LOR^{(i)}) = \text{Var}(lpc) = 1/15 + 1/85,$$

as $LOR^{(i)}$ is known (i.e., fixed on every inner cycle k) and therefore has zero variance. Using this form of calculation with 100,000 simulations, EVPPI for LOR alone was $3920.

EVSI for the LOR

What would be the EVSI for a study allocating $n_T = 200$ patients to new treatment and another $n_C = 200$ to standard care? The procedure here is to set all parameters except p_T and LOR to their prior expectations and then (step B1(a)) draw samples $p_C^{(i)}$ from the prior distribution of $p_C \sim$ beta(15,85) and a prior $LOR^{(i)}$ from its prior distribution $N(-1.5,3)$ (step B1(a)). On step B1(b), these 2 samples are combined to create a sample $p_T^{(i)} = \text{logit}(p_C^{(i)}) + LOR^{(i)}$.

Then (step B2(a)), draw binomial samples to serve as numerators for each arm in the proposed new study:

$$r_C^{(i)} \sim \text{Bin}(p_C^{(i)}, n_C) \text{ and } r_T^{(i)} \sim \text{Bin}(p_T^{(i)}, n_T).$$

Next (step B2(b)), the new data on simulation i are formed into an LOR $\mu_D^{(i)}$ with precision $\tau_D^{(i)}$. Whenever $r_C^{(i)}$ or $r_T^{(i)} = 0$ or n_C, n_T, the standard zero cell correction is carried out, adding 0.5 to each numerator and 1 to each denominator for that cycle. Using standard formulae, we then define the sample LOR, $\mu_D^{(i)}$ and its precision $\tau_D^{(i)}$.

We then (step B3(a)) obtain a posterior mean and precision for $LOR \mid D^{(i)}$ as a precision weighted average of $\mu_D^{(i)}$ and the prior mean –1.50. The posterior precision is $3 + \tau_D^{(i)}$. Finally, we convert our information on the posterior mean LOR with information on the prior mean of p_C to obtain a mean and variance for the posterior of $p_T \mid D^{(i)}$:

$$lp_T^{(i)} = E\,[\text{logit}(p_T \mid D^{(i)})] = \log(14.5/84.5) + E(LOR \mid D^{(i)})$$

and then use the Taylor series formulae in Table 2 to obtain a posterior mean for $p_T \mid D^{(i)}$:

$$h^{(i)} = \exp(lp_T^{(i)})/(1 + \exp(lp_T^{(i)}))$$

$$E[p_T \mid D^{(i)}] = h^{(i)} + h^{(i)}(1 - h^{(i)})(1 - 2h^{(i)})([1/(3 + \tau_D^{(i)})] + 1/15 + 1/85)/2.$$

Note that the variance expression in the Taylor series formula reflects the fact that the posterior variance of $\text{logit}(p_T \mid D^{(i)})$ is the sum of the prior variance of $\text{logit}(p_C)$ and the posterior variance of $LOR \mid D^{(i)}$. The net benefit equations are then run with the posterior mean of p_T and all other parameters, including p_C, taking their prior mean values. An allocation of $n = 200$ to each arm gives an EVSI of $3260.

EVSI for a Randomized Clinical Trial (RCT) Gaining Information on LOR, p_{SE} and Q_E

An RCT to compare new treatment and standard care is also an opportunity to obtain more information on p_{SE} in patients allocated to the new treatment and more information on the quality-adjusted life year (QALY) decrement Q_E that follows the critical event. Consider a design, for example, in which the investigator is able to choose an allocation for n_C and n_T, with the aim of gaining information on the LOR but also to use the n_T patients allocated to the new treatment to gain more information on the probability of side effects p_{SE}. It is also proposed to use the patients experiencing the critical event in both trial arms to contribute further information on the QALY decrement Q_E. Note that the number available for this is not known in advance. The simulation approach proposed can readily accommodate this, as the calculation of $E[Q_E \mid D^{(i)}]$ on each cycle i can be based on the sample size $n^{(i)} = r_C^{(i)} + r_T^{(i)}$, the total number of patients suffering the critical event in both arms of the proposed RCT, even though this varies across Monte Carlo cycles. The net benefit equations have the form

$$B(t, E\,[p_{SE} \mid D^{(i)}], E[Q_E \mid D^{(i)}], 0.1499, E\,[p_T \mid D^{(i)}]).$$

Averaging over all cycles, the EVSI for an RCT with 200 allocated to each arm under these circumstances is $8330. This type of calculation, in which the same study provides information on several model parameters, demonstrates the flexibility of the simulation-based approach to EVSI in the context of optimal trial design.

REFERENCES

1. Stinnett A, Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analyses. Med Decis Making. 1998;18:S68–S80.
2. Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of heath care technologies. J Health Econ. 1999;18:341–64.
3. Schlaiffer R. Probability and Statistics for Business Decisions. New York: McGraw-Hill; 1958.
4. Raiffa H, Schlaiffer R. Applied Statistical Decision Theory. New York: Wiley Interscience; 1967.
5. Finkel AS, Evans JS. Evaluating the benefits of uncertainty reduction in environmental health risk management. J Air Pollut Control Assoc. 1987;37:1164–71.
6. Reichard EG, Evans JS. Assessing the value of hydrogeologic information for risk-based remedial action decisions. Water Resour Res. 1989;25:1451–60.
7. Taylor AC, Evans JS, McKone TE. The value of animal test information in environmental control decisions. Risk Anal. 1993;13:403–12.
8. Thompson KM, Evans JS. The value of improved national exposure information for perchloroethylene (Perc): a case study for dry cleaners. Risk Anal. 1997;17:253–71.
9. Felli JC, Hazen GB. Sensitivity analysis and the expected value of perfect information. Med Decis Making. 1998;18:95–109.
10. Fenwick E, Claxton K, Sculpher M, Briggs A. Improving the Efficiency and Relevance of Heath Technology Assessment: The Role of Iterative Decision Analytic Modelling. Report no. 179. York, United Kingdom: Centre for Health Economics, University of York; 2000.

11. Claxton K, Posnett J. An economic approach to clinical trial design and research priority-setting. Health Econ. 1996;5:513–24.
12. Davis DR, Kisiel CC, Duckstein L. Bayesian decision theory applied to design in hydrology. Water Resour J. 1972;8:33–41.
13. Wagner BJ. Sampling design methods for groundwater modelling under uncertainty. Water Resour J. 1995;31:2581–91.
14. Dakins ME, Toll JE, Small MJ, Brand KP. Risk-based environmental remediation: Bayesian Monte Carlo analysis and the expected value of sample information. Risk Anal. 1996;16:67–79.
15. Claxton K, Lacey LF, Walker SG. Selecting treatments: a decision theoretic approach. Journal of the Royal Statistical Society (A). 2000;163:211–26.
16. Phillips CV. The economics of more research is needed. Int J Epidemiol. 2001;30:771–6.
17. Sterne JAC, Davey SG. Sifting the evidence—what's wrong with significance tests. BMJ. 2001;322:226–31.
18. Peto R, Baigent C. Trials: the next 50 years. Large scale randomised evidence of moderate benefits. BMJ. 1998;317:1170.
19. Claxton K. Bayesian approaches to the value of information: implications for the regulation of new pharmaceuticals. Health Econ. 1999;8:269–74.
20. Doubilet P, Begg CB, Weinstein MC, Braun P, McNeill BJ. Probabilistic sensitivity analysis using Monte Carlo simulation: a practical approach. Med Decis Making. 1985;5:157–77.
21. Critchfield GC, Willard KE. Probabilistic analysis of decision trees using Monte Carlo simulation. Med Decis Making. 1986;6:85–92.
22. Thompson KM, Burmaster DE, Crouch EAC. Monte Carlo techniques for quantitative uncertainty analysis in public health risk assessments. Risk Anal. 1992;12:53–63.
23. Manning WG, Fryback DG, Weinstein MC. Reflecting uncertainty in decision analysis. In: Gold MR, Siegel JE, Russell LB, Weinstein MC, eds. Cost-Effectiveness in Health and Medicine. New York: Oxford University Press; 1996. p 247–75.
24. Commonwealth Department of Health, Housing and Community Services. Guidelines for the Pharmaceutical Industry on Preparation of Submissions to the Pharmaceutical Benefits Advisory Committee. Canberra, Australia: APGS; 1992.
25. Ontario Ministry of Health. Ontario Guidelines for Economic Analyses of Pharmaceutical Products. Ontario (Canada): Ministry of Health; 1994.
26. Langley PC, Donaldson C, eds. Formulary submission guidelines for the Blue Cross and Blue Shield of Colorado and Nevada. Structure, application and manufacturer responsibilities. Pharmacoeconomics. 1999;16:211–24.
27. Claxton K, Sculpher NJS, Drummond M. A rational framework for decision making by the National Institute for Clinical Excellence. Lancet. 2002;360:711–5.
28. National Institute for Clinical Excellence. Guidance for manufacturers and sponsors. 2001. Available from: URL: www.nice.org.uk
29. DerSimonian R, Laird N. Meta-analysis of clinical trials. Control Clin Trials. 1986;7:177–88.
30. Thompson SG. Why sources of heterogeneity in meta-analyses should be investigated. BMJ. 1994;309:1351–5.
31. Lau J, Ioannidis JPA, Schmid CH. Summing up evidence: one answer is not always enough. Lancet. 1998;351:123–7.
32. Briggs AH. A Bayesian approach to stochastic cost-effectiveness analysis. Health Econ. 1999;8:257–61.
33. Heitjan DF. Fieller's method and net health benefits. Health Econ. 2000;9:327–35.
34. Laska EM, Meisner M, Siegel C, Wanderling J. Statistical cost-effectiveness analysis of two treatments based on net health benefits. Stat Med. 2001;20:1279–302.
35. O'Hagan A, Stevens JW. A framework for cost-effectiveness analysis from clinical trial data. Health Econ. 2001;10:303–15.
36. Claxton K, Thompson KM. A dynamic programming approach to efficient clinical trial design. J Health Econ. 2001;20:797–822.
37. Bernardo JM, Smith AFM. Bayesian Theory. New York: Wiley; 1994.
38. O'Hagan A. Kendall's Advanced Theory of Statistics. Volume 2B, Bayesian Inference. London: Edward Arnold; 1994.
39. Hazen GB. Factored stochastic trees: a tool for solving complex temporal medical decision models. Med Decis Making. 1993;13:227–36.
40. Hazen GB, Pellissier JM. Recursive utility for stochastic trees. Oper Res. 1996;44:788–809.
41. Hazen GB, Pellissier JM, Sounderpandian J. Stochastic-tree models in medical decision making. Interfaces. 1998;28:64–80.
42. Hazen G. Preference factoring for stochastic trees. Manage Sci. 2000;46:389–403.
43. Beck JR, Pauker S. The Markov process in medical prognosis. Med Decis Making. 1983;3:419–58.
44. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. Health Econ. 1983;13:322–38.
45. Grimmett GDR, Stirzaker DR. Probability and Random Processes. 2nd ed. Oxford (UK): Oxford University Press; 1992.
46. Shadish WR, Haddock CK. Combining stimates of effect size. In: Cooper H, Hedges LV, eds. The Handbook of Research Synthesis. New York: Russell Sage Foundation; 1994. p 261–81.
47. Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F. Systematic reviews of trials and other studies. Health Technol Assess. 1998;2:1–276.
48. Gelman A, Carlin JG, Stern HS, Rubin DB. Bayesian Data Analysis. London: Chapman and Hall; 1995.
49. Deeks JJ. Issues on the selection of a summary statistic for meta-analysis of clinical trials with binary outcomes. Stat Med. 2002;21:1575–600.
50. Rubin DB. A new perspective. In: Wachter KW, Straf ML, eds. The Future of Meta-analysis. New York: Russell Sage Foundation; 1990.
51. Berkey CS, Hoaglin DC, Mosteller F, Colditz GA. A random effects regression model for meta-analysis. Stat Med. 1995;14:395–411.
52. Thompson SG, Smith TC, Sharp SJ. Investigating underlying risk as a source of heterogeneity in meta-analysis. Stat Med. 1997;16:2741–758.
53. Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. Stat Med. 1999;18:2693–708.
54. O'Hagan A, Kennedy MC, Oakley JE. Uncertainty analysis and other inference tools for complex computer codes. In: Bernardo JM, Berger JO, Dawid AP, Smith AFM, eds. Bayesian Statistics 6. Oxford (UK): Clarendon; 1998. p 503–24.
55. Rasmussen CE, Ghahramani Z. Bayesian Monte Carlo. 2002. Available from: URL: http://www-2.cs.cmu.edu/Groups/NIPS/NIPS2002/NIPS2002preproceedings/papers/revised/AA01-draft. pdf
56. Oakley J, O'Hagan A. Probabilistic Sensitivity Analysis of Complex Models: A Bayesian Approach. Report no. 525/02. Department of Probability and Statistics, University of Sheffield; 2002.
57. Stevenson MD, Oakley J, Chilcott JB. 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. DEEMS Workshop Synthesizing the Evidence for Economic Evaluation; 2002 Apr 11–12; Oxford, UK.
58. Eddy DM, Hasselblad V, Shachter R. Meta-analysis by the Confidence Profile Method. London: Academic Press; 1992.
59. Dilks DW, Canale RP, Meier PG. Development of Bayesian Monte Carlo techniques for water quality model uncertainty. Ecol Model. 1992;62:149–62.
60. Brand KP, Small MJ. Updating uncertainty in an integrated risk assessment: conceptual framework and methods. Risk Anal. 1995;15:719–31.

61. Sohn MD, Small MJ, Pantazidou M. Reducing uncertainty in site characterisation using Bayes Monte Carlo methods. Journal of Environmental Engineering. 2000;126:893–902.
62. Ades AE, Cliffe S. Markov Chain Monte Carlo estimation of a multi-parameter decision model: consistency of evidence and the accurate assessment of uncertainty. Med Decis Making. 2002; 22:359–71.
63. Carlin JG, Louis TA. Bayes and Empirical Bayes Methods for Data Analysis. 2nd ed. Boca Raton (FL): Chapman & Hall/CRC; 2000.
64. Smith TC, Spiegelhalter DJ, Thomas SL. Bayesian approaches to random-effects meta-analysis: a comparative study. Stat Med. 1995;14:2685–99.
65. Warn DE, Thompson SG, Spiegelhalter DJ. Bayesian random effects meta-analysis of trials with binary outcomes: method for absolute risk difference and relative risk scales. Stat Med. 2002;21:1601–23.
66. Abrams K, Sanso B. Approximate Bayesian inference for random effects meta-analysis. Stat Med. 1998;17:201–18.
67. Claxton K, Ades AE. Efficient research design: an application of value of information analysis to an economic model of zanamivir. Proceedings of the 24th Meeting of the Society for Medical Decision Making; 2002 Oct 20–23; Baltimore, MD.
68. Johnson NL, Kotz S, Kemp AW. Distributions in Statistics: Discrete Distributions. New York: Wiley; 1992.