



HHS Public Access

Author manuscript

Med Decis Making. Author manuscript; available in PMC 2021 November 22.

Published in final edited form as:

Med Decis Making. 2018 February ; 38(2): 174–188. doi:10.1177/0272989X17715627.

A Gaussian Approximation Approach for Value of Information Analysis

Hawre Jalal,

Department of Health Policy and Management, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA

Fernando Alarid-Escudero

Department of Health Policy and Management, School of Public Health, University of Minnesota, Minneapolis, MN, USA

Abstract

Most decisions are associated with uncertainty. Value of information (VOI) analysis quantifies the opportunity loss associated with choosing a suboptimal intervention based on current imperfect information. VOI can inform the value of collecting additional information, resource allocation, research prioritization, and future research designs. However, in practice, VOI remains underused due to many conceptual and computational challenges associated with its application. Expected value of sample information (EVSI) is rooted in Bayesian statistical decision theory and measures the value of information from a finite sample. The past few years have witnessed a dramatic growth in computationally efficient methods to calculate EVSI, including metamodeling. However, little research has been done to simplify the experimental data collection step inherent to all EVSI computations, especially for correlated model parameters. This article proposes a general Gaussian approximation (GA) of the traditional Bayesian updating approach based on the original work by Raiffa and Schlaifer to compute EVSI. The proposed approach uses a single probabilistic sensitivity analysis (PSA) data set and involves 2 steps: 1) a linear metamodel step to compute the EVSI on the preposterior distributions and 2) a GA step to compute the preposterior distribution of the parameters of interest. The proposed approach is efficient and can be applied for a wide range of data collection designs involving multiple non-Gaussian parameters and unbalanced study designs. Our approach is particularly useful when the parameters of an economic evaluation are correlated or interact.

Keywords

expected value of sample information; metamodeling; probabilistic sensitivity analysis; uncertainty; value of information analysis

Reprints and permissions: sagepub.com/journalsPermissions.nav

Corresponding Author: Hawre Jalal, Department of Health Policy and Management, Graduate School of Public Health, University of Pittsburgh, 703 Parran Hall, 130 Desoto St, Pittsburgh, PA 15261, USA. (hjalal@pitt.edu).

Supplementary Material

Supplementary material for this article is available on the *Medical Decision Making* Web site at <http://journals.sagepub.com/home/mdm>.

Economic evaluation is a framework to quantify the expected costs and benefits across several strategies and determine the overall optimal strategy by simplifying real-life complexity. However, recommendations from these analyses are rarely definitive due to uncertain input parameters that may translate to “suboptimal” recommendations. Choosing the optimal strategy at the *time of decision making* can incur a large opportunity cost because the overall optimal strategy might be suboptimal in certain realizations of the input parameter values. Value of information analysis (VOI) quantifies this opportunity cost and sets the maximum limit on how much new information is potentially worth. For example, expected value of perfect information (EVPI) measures the value of eliminating uncertainty from all model parameters while expected value of partial perfect information (EVPPI) measures the EVPI for a subset of parameters of interest. Eliminating uncertainty can only be achieved with an infinitely large sample, which is not feasible in practice. Thus, expected value of sample information (EVSI) is generally more informative than EVPPI because it measures the value of reducing uncertainty from a finite sample of size n .

EVSI is defined as the value of information from a *finite* sample (n) for 1 or more parameters in a decision-analytic model.¹ EVSI is rooted in Bayesian statistical decision theory and measures the value of information by generating potential new data before any actual data collection is conducted.²⁻⁵

Thus, EVSI can be of important practical application in research prioritization, allocation of scarce resources, and informing the type and size of the study design for future data collection efforts.^{1,6-8} In practice, however, EVSI remains underused due to many computational and conceptual challenges associated with its practical implementation.⁹

EVSI is currently an area of active methodological research. In the past 2 decades, several new approaches have been proposed to help researchers implement EVSI. For example, Brennan and others¹⁰ proposed a nested 2-level Monte-Carlo simulation (2MCS) to numerically compute EVSI as an approximation to the double integrals involved in the EVSI equation. Although their approach applies to many situations, in practice it is often computationally expensive. In addition, it may be challenging to implement when there exist correlations between the parameters of interest (θ_I) and the complementary set of the parameters (θ_C), where $\theta = \{\theta_I, \theta_C\}$. Brennan and Kharroubi^{5,11} have proposed a Bayesian Laplace approximation that replaces both the Bayesian updating and the inner Monte Carlo sampling to compute the posterior expectation of the parameters of interest. However, this approach may still be computationally expensive and generally assumes that the observations are independent and identically distributed (iid), where the number of study participants represents the sample size used to estimate the mean and standard error of the parameters. Ades and others¹² have developed an analytic approach to calculate EVSI that is generally computationally efficient. However, this technique is limited to models with certain structural forms and certain independence assumptions on the parameters are made.

More recent developments have used a metamodeling approach to simplify the inner expectation. For example, Jalal and others¹³ adopted the unit normal integral function (UNLI) in addition to linear regression metamodeling to calculate EVSI. Their technique assumes that the parameters of interest have a normal distribution, and it can only be used

for a single parameter at a time if there are multiple strategies. Strong and others¹⁴ have proposed a generalized additive models (GAM) metamodeling approach to approximate the inner expectation by defining a summary statistic. Furthermore, Menzies¹⁵ adopts an efficient algorithm that uses importance sampling in addition to a metamodel to simplify EVSI computations. This approach requires the user to define 2 functions: a function for creating data sets from the prior and another function for computing the likelihood given the data sets. Thus, significant advances have been made in increasing the computational efficiency of EVSI.¹⁶ However, little research has been done to overcome a second challenge of EVSI, which involves Bayesian updating for study designs that involve multiple correlated parameters or when a study design is unbalanced involving several parameters informed by different sample sizes.

In this study, we propose a simplification of the data collection process via a Gaussian approximation (GA) of the traditional Bayesian updating process. This approximation generalizes Raiffa and Schlaifer's original work^{3,4,17} for a normal prior and normal data likelihood to a wide range of univariate and multivariate non-Gaussian distributions. The GA approach only requires a data set of prior values and their associated prior effective sample size n_0 . In addition, the GA allows complex correlation structures among the model parameters θ . Similar to the previous approaches,¹³⁻¹⁵ our approach uses a metamodel on a single probabilistic sensitivity analysis (PSA) data set, which is routinely obtained in economic evaluations.^{18,19}

Methods

EVSI

EVSI is typically expressed as^{10,12}

$$EVSI_{\theta_I} = \mathbb{E}_X \left[\max_d \mathbb{E}_\theta [B(d, \theta) | X] \right] - \max_d \mathbb{E}_\theta [B(d, \theta)], \quad (1)$$

where θ are the parameters of the decision model and consist of the parameters of interest θ_I and the complementary parameters θ_C , X are new potential data, and $B(d, \theta)$ is the net benefit from strategy d . Thus, EVSI is the difference between the expected maximum benefit given *new* potential data X and the maximum benefit with current information. Alternatively, EVSI can be expressed as a function of the opportunity loss from choosing a suboptimal decision rather than difference between 2 benefits, such that

$$EVSI_{\theta_I} = \mathbb{E}_X \max_d \mathbb{E}_\theta [L(d, \theta) | X], \quad (2)$$

where $L(d, \theta) = B(d, \theta) - B(d^*, \theta)$ is the opportunity loss from choosing the optimal decision d^* among all possible decisions D , given a set of parameter values θ . (Appendix A shows the steps of deriving equation (2) from equation (1).)

The inner expectation $\mathbb{E}_\theta [L(d, \theta) | X]$ computes the opportunity loss from choosing d instead of d^* given a set of parameter values θ and a potential candidate for X . Here, we are interested in avoiding the maximum loss from any strategy d , and therefore we take the

$\max_d \mathbb{E}_\theta [L(d, \theta) | X]$. Since we do not know what the new data X may look like a priori, we take the expectation over all possible candidates of X and compute $\mathbb{E}_X \max_d \mathbb{E}_\theta [L(d, \theta) | X]$. Thus, EVSI sets our upper limit on how much we should be willing to pay to avoid the expected maximum opportunity loss over all strategies.

The 2 nested expectations in equations (1) and (2) could be approximated with 2 nested Monte Carlo loops,¹⁰ which are often computationally expensive even for relatively simple models. We simplify equation (2) using 2 approximations: 1) we approximate the opportunity losses L as a linear relation of the parameters of interest (or a transformation of these parameters), and 2) we propose a Gaussian approximation approach to further simplify the Bayesian updating process to compute $\theta | X$. Previous studies have described the first approximation.^{13,14} The main contribution of this article is related to simplifying some of the computational and conceptual burdens of conducting the Bayesian updating and the data generation process using a Gaussian approximation approach.

Linear Metamodel Approximation

We propose a linear metamodeling approach to simplify the inner expectation in equation (2). The linear metamodel (LM) is a second model involving regressing the model output, such as the opportunity loss L , on the input parameters (or a transformation of these parameters).²⁰ LM has been used on PSA data sets to reveal the characteristics of a model²¹ and to simplify VOI computations.^{13,14,22} The LM can range from a simple linear regression that describes a linear relation between the model inputs and outputs to spline regressions that are more flexible and can describe a linear relation between transformations of a set of basis functions of θ and L .²³

In VOI analysis, we are generally interested in the opportunity loss $L(d, \theta)$ explained by θ_j ; therefore, we can define a linear metamodel as

$$L(d, \theta) = \beta_0 + \beta_1 \theta_j + e, \quad (3)$$

where β_0 is an intercept, β_1 is a regression coefficient for θ_j , and e is a residual term. This residual term captures the complementary parameters, θ_C , and the nonlinearity in the model. In addition, e also captures Monte Carlo noise in stochastic models, such as microsimulation models. Thus, the expected loss is estimated using the LM given θ_j by

$$\hat{L}(d, \theta) | \theta_j = \hat{\beta}_0 + \hat{\beta}_1 \theta_j, \quad (4)$$

where $\hat{\beta}_0$ and $\hat{\beta}_1$ can be estimated by regressing L on θ_j in equation (3) via ordinary least squares (OLS) using a PSA data set. Equation (4) simplifies the inner expectation in EVPPI because it computes the expected opportunity loss while averaging over θ_C . Various versions of this equation have been used as a step in the calculation of EVPPI.^{13,22} This simple equation is appealing because the regression coefficients can be easily estimated in any standard statistical software such as using the function `lm` in R by defining `lm1 <- lm (L ~ theta_I)`. Then, it is usually straightforward to compute the conditional loss using the `predict` function, such that `L_hat <- predict(lm1)`.

In EVSI, the inner expectation measures the opportunity loss given the updated distribution of θ after observing a potential data set $X^{(i)}$ from all possible candidates. This inner expectation then needs to be reevaluated for all the possible $X^{(i)}$ candidates that are sampled from the outer expectation. This nested expectation structure renders EVSI computationally expensive via Monte Carlo simulation because it requires the model to be reevaluated many times inside these nested expectations to compute $L(d, \theta)$ for each sample from the posterior distribution of $\theta_I|X^{(i)}$ and the prior distribution of θ_C .

The LM is a linear approximation of the original simulation model. This LM defines a linear relation between $L(d, \theta)$ and θ . That is, the structure of the LM is independent on the values of θ once the LM is defined for a particular model. Thus, we can plug in the posterior mean of θ_I , $\mathbb{E}_\theta[\theta_I|X^{(i)}]$ and use the LM to compute the conditional loss given each data set $X^{(i)}$. The LM simplifies the nested computation in EVSI because the inner expectation can be simplified to a linear combination of the posterior expectation of θ_I . Specifically, the inner expectation of EVSI given $X^{(i)}$ in equation (2), $\mathbb{E}_\theta[L(d, \theta) | X^{(i)}]$, can be estimated as a linear function of the posterior mean of θ_I , $\mathbb{E}_\theta[\theta_I|X^{(i)}]$, such that

$$\tilde{L}_d^{(i)} = \hat{\beta}_0 + \hat{\beta}_1 \phi^{(i)}, \tag{5}$$

where we denote $\tilde{L}_d^{(i)} = \hat{L}(d, \theta)|X^{(i)}$, and $\phi^{(i)} = \mathbb{E}_\theta[\theta_I|X^{(i)}]$.

Notice that if θ_I and θ_C are correlated a priori, then data collection on θ_I will likely inform θ_C as well. Fortunately, $\hat{\beta}_1$ captures this correlation, and it allows the opportunity loss to also reflect the impact of the new data X on θ_C .

Thus, we only need $\hat{\beta}_0$, $\hat{\beta}_1$, and ϕ to compute \tilde{L}_d for each strategy. In R, we can achieve this by using the predict function again using ϕ instead, such that `L_tilde <- predict(lm1, newdata = phi)`. Thus, we can approximate EVSI using

$$\widehat{\text{EVSI}}_{\theta_I} = \mathbb{E}_X \max_d(\tilde{L}_d), \tag{6}$$

where all the variables have been previously defined.

So far, we have assumed that we know ϕ , θ_I is a single parameter, and the relation between L and θ_I is linear. In the next section, we show how to approximate ϕ using our Gaussian approximation approach, and in the following sections, we relax the other assumptions.

Gaussian Approximation of ϕ

The posterior mean ϕ is a random variable because X is random. The distribution of ϕ is often referred to as the “preposterior” distribution because it defines the distribution of the posterior mean prior to actual data collection. In this section, we first show that if θ_I has a Gaussian distribution and the data likelihood of $X|\theta_I$ is also Gaussian, then we can estimate ϕ directly from θ_I by assuming that the prior is the only source of randomness in ϕ . Later,

we apply this relation to non-Gaussian prior-likelihood pairs and compare the approximation numerically to traditional Bayesian updating for these non-Gaussian cases.

Let's assume that our prior knowledge of θ_I is based on a study of size n_0 , such that

$$\theta_I \sim N\left(\mu_0, \frac{\sigma^2}{n_0}\right), \quad (7)$$

where μ_0 and σ^2 are the population mean and variance of θ_I , respectively.

We are interested in generating a “new” sample mean from n new participants given the distribution of θ_I . This new mean X also follows a Gaussian distribution such that

$$X|\theta_I \sim N\left(\theta_I, \frac{\sigma^2}{n}\right). \quad (8)$$

Because the data likelihood and the prior are conjugate, we can express the posterior distribution as

$$\theta_I|X \sim N\left(\frac{n_0\mu_0 + nX}{n_0 + n}, \frac{\sigma^2}{n + n_0}\right). \quad (9)$$

The mean of this distribution is ϕ :

$$\phi = \frac{n_0\mu_0 + nX}{n_0 + n}. \quad (10)$$

Notice that ϕ is a function of the marginalized X , and as such it is also a random variable that has the same mean as θ_I , that is,

$$\begin{aligned} \mathbb{E}_X[\phi] &= \mathbb{E}_X\left[\mathbb{E}_{\theta_I}[\theta_I | X]\right] \\ \mathbb{E}_X[\phi] &= \mathbb{E}_{\theta_I}[\theta_I] \\ \mathbb{E}_X[\phi] &= \mu_0. \end{aligned} \quad (11)$$

It is important here to distinguish between the *posterior* mean, which is a random variable denoted by ϕ , from the mean of ϕ with respect to X , which is a single value equal to μ_0 . The variance of ϕ can be defined from both equation (10) and the variance of the marginalized X , which is equal to the sum of the prior variance and the variance of the data likelihood,²⁴ such that $\text{Var}(X) = \sigma^2/n + \sigma^2/n_0$. Therefore, the variance of ϕ is expressed as

$$\begin{aligned} \text{Var}_X(\phi) &= \left(\frac{n}{n_0+n}\right)^2 \left(\frac{\sigma^2}{n} + \frac{\sigma^2}{n_0}\right) \\ \text{Var}_X(\phi) &= \frac{n}{n_0+n} \left(\frac{\sigma^2}{n_0}\right) \\ \text{Var}_X(\phi) &= v \text{Var}(\theta_I). \end{aligned} \tag{12}$$

The term v is defined as the variance fraction $n/(n_0 + n)$, indicating that the variance of the preposterior distribution is always a fraction of the prior variance as it has been shown previously.^{3,12} The $\text{Var}_X(\phi)$ tends to zero as $n \rightarrow 0$ because with a small sample, we know that our posterior mean will be very similar to the prior mean. Conversely, $\text{Var}_X(\phi) \rightarrow \text{Var}(\theta_I)$ as $n \rightarrow \infty$ because the prior is the source of all information; thus, as n increases, our uncertainty regarding ϕ will mimic prior uncertainty.¹² In addition, the expectation of ϕ equals μ_0 regardless of n . Figure 1 illustrates the relation between the prior and the preposterior distributions for various n in the Gaussian-Gaussian prior-likelihood case.

Equations (11) and (12) define the mean and variance of ϕ , respectively. From equation (10), we can see that ϕ is normally distributed because X is normally distributed and the rest of the terms are constants. As a result, we can express ϕ in terms of its mean, variance, and a standard normal distribution $Z \sim N(0, 1)$, such that

$$\phi = \mu_0 + Z \frac{\sqrt{v}\sigma}{\sqrt{n_0}}. \tag{13}$$

Adding and subtracting $\sqrt{v}\mu_0$ results in

$$\phi = \mu_0 + Z \frac{\sqrt{v}\sigma}{\sqrt{n_0}} + \sqrt{v}\mu_0 - \sqrt{v}\mu_0, \tag{14}$$

and rearranging the terms produces

$$\phi = \sqrt{v} \left(\mu_0 + Z \frac{\sigma}{\sqrt{n_0}} \right) + (1 - \sqrt{v})\mu_0. \tag{15}$$

Since X is not observed, and θ_I is the source of all information, we attribute all randomness in ϕ to randomness in the prior of θ_I , which allows us to replace $\mu_0 + Z\sigma/\sqrt{n_0}$ with θ_I and obtain

$$\phi = \sqrt{v}\theta_I + (1 - \sqrt{v})\mu_0. \tag{16}$$

Again, if the proposed data collection sample relative to n_0 is small, $\sqrt{v} \rightarrow 0$ and the estimated posterior mean will be very similar to the prior mean (i.e., $\phi \rightarrow \mu_0$). $\phi \rightarrow \mu_0$). But

if the proposed sample size is relatively large, $\sqrt{v} \rightarrow 1$ the posterior mean converges to the prior (i.e., $\phi \rightarrow \theta_j$).

Equation (16) can be extended to the multivariate Gaussian prior Gaussian likelihood because it can be applied to each component of the multivariate Gaussian distribution. The resulting correlation structure of these preposterior components will then be the same as the correlation structure among the prior components. This equation plays a key role in our approach to compute EVSI because it simplifies Bayesian updating by computing the posterior mean as a simple weighted average of the prior and the prior mean. This relation allows us to apply equation (16) in other non-Gaussian prior-likelihood pairs. Later, we test the performance of this approximation in a variety of non-Gaussian prior-likelihood distribution pairs where the priors and the data likelihoods are conjugate or nonconjugate and univariate or multivariate.

Using equation (16), a researcher only needs a set of K PSA simulations to compute ϕ . Thus, given an observation $\theta_j^{(i)}$, an estimate $\bar{\theta}_j$ of μ_0 and $v = n/(n + n_0)$, we can estimate $\phi^{(i)}$ such that

$$\phi^{(i)} = \sqrt{v}\theta_j^{(i)} + (1 - \sqrt{v})\bar{\theta}_j, \quad (17)$$

for all $i = 1, \dots, K$, where $\bar{\theta}_j$ is the prior mean and can be estimated from the same PSA sample $\bar{\theta}_j = \frac{1}{K} \sum_{i=1}^K \theta_j^{(i)}$.

EVSI with Gaussian Approximation and Linear Metamodeling

After obtaining ϕ and estimating $\hat{\beta}_0$ and $\hat{\beta}_1$, we can compute $\tilde{L}_d^{(i)}$ from equation (5) and approximate EVSI by

$$\widehat{\text{EVSI}}_{\theta_j} = \frac{1}{K} \sum_{i=1}^K \max_d(\tilde{L}_d^{(i)}). \quad (18)$$

In the previous equation, \tilde{L}_d is an approximation of the loss as a linear function of a single parameter. In the next couple sections, we extend the definition of \tilde{L}_d to cases with more than one parameter and where the relation between the parameters and the loss function can be expressed with more flexible relations, such as splines.

Splines as Linear Metamodels

If the relation between L and θ_j is not linear, it is still possible to use a linear metamodel by defining a linear relation between L and a transformation of θ_j , namely, $f(\theta_j)$. A common choice for $f(\theta_j)$ is a spline function, such that $f(\theta_j) = a_0 + ba_1$, where b represents a set of basis functions of θ_j that form the spline, a_0 is the new intercept, and a_1 is a vector of regression coefficients of the set of basis functions, b .²³

We now can reexpress equation (3) in terms of splines as

$$L(d, \theta) = \alpha_0 + b\alpha_1 + u, \quad (19)$$

where u is the new residual term. Similar to the `lm` function above, we can use the `gam` function in the R package `mgcv` to estimate the coefficients α_0 and α_1 by transforming θ_I into a set of basis functions, such that `lm2 <- gam(L ~ s(theta_I))`.²³ Then, we can use the fitted values using `lm2$fitted` to obtain the loss L conditional on θ_I . We refer to this conditional loss as \hat{L} .

Using equation (19), we can also estimate the loss conditional on ϕ , referred to as \hat{L} . Because equations (16) and (19) are both linear, we can apply the variance reduction to the basis functions directly by using the `predict.ga` function provided in the appendix to estimate \tilde{L} , such that `L_tilde <- predict.ga(lm2, n, n0)`. Finally, we can use equation (18) to compute EVSI.

Notice that by letting $n \rightarrow \infty$, $\tilde{L}_d^{(i)} \rightarrow \hat{L}_d^{(i)}$, transforming equation (18) to the equation of EVPPI as a function of the loss conditional on the prior θ_I

$$\widehat{\text{EVPPI}}_{\theta_I} = \frac{1}{K} \sum_{i=1}^K \max_d(\hat{L}_d^{(i)}). \quad (20)$$

In Box 1 and Appendix B, we provide a step-by-step summary and examples of R code to compute EVPPI and EVSI with Gaussian approximation and spline linear metamodeling.

EVSI for Multiple Parameters

Here we consider the case where $\theta_I = \{\theta_{I1}, \theta_{I2}\}$, and the loss function can be expressed as

$$L(d, \theta) = \delta_0 + f_1(\theta_{I1}) + f_2(\theta_{I2}) + f_{12}(\theta_{I1}, \theta_{I2}) + \epsilon, \quad (21)$$

where δ_0 is the new intercept; f_1 , f_2 , and f_{12} represent the spline functions for θ_{I1} , θ_{I2} , and their interaction, respectively; and ϵ is the new residual term. Equation (21) is often referred to as a GAM and can be estimated using `mgcv`.²³

This setup also allows one to compute EVSI for studies that collect information on multiple parameters with different n_0 . In addition, this setup can be applied to unbalanced data collection study designs that propose to collect information on more than 1 parameter from a different number of participants. These data collection designs can be implemented by computing different v for various parameters of interest (e.g., θ_{I1} and θ_{I2}). Appendix B provides further R code implementations for the Gaussian approximation when there are multiple parameters and unbalanced data collection designs where n and n_0 vary by the parameter of interest.

Estimating n_0

The effective sample size n_0 denotes the “amount” of information in the prior.²⁵ In some instances, n_0 can be directly inferred from the prior distribution of the parameters and the

data likelihood or elicited from experts' judgments. Table 1 includes examples of direct assessment of n_0 from 3 different prior distributions. For example, n_0 for a beta prior $p \sim \text{beta}(\alpha, \beta)$ and binomial data likelihood $x|p \sim \text{binomial}(p, n)$ can be computed as $n_0 = \alpha + \beta$, where n_0 is simply the sum of the number of successes and failures. It is important to emphasize that the data likelihood also determines n_0 . For example, in Table 1, we illustrate the case of a gamma prior and exponential data likelihood, in which $n_0 = a$, where a represents the shape parameter of the gamma distribution. However, if the data likelihood has a Poisson distribution to count the number of events, $n_0 = 1/b$, where b represents the scale parameter of the gamma distribution. We refer the reader to Morita and others²⁵ for a more detailed discussion on computing n_0 .

Whenever n_0 is not readily available, we propose 2 alternative approaches to estimate it, both using the Gaussian approximation. The first approach involves computing a summary statistic, and the second approach involves estimating n_0 indirectly from a Bayesian data collection experiment. The second approach can be generalized to a wide variety of situations, including cases in which the prior and data likelihoods are not conjugate.

Computing n_0 using a summary statistic S .—This Gaussian approximation approach to compute n_0 requires a summary statistic S to describe the data collection step. In the Gaussian prior-Gaussian likelihood case, the marginalized X represents the summary statistic, that is, $S = X$. Therefore, $\text{Var}(S) = \sigma^2/n + \sigma^2/n_0$. Since, we know $\text{Var}(\theta_I) = \sigma^2/n_0$, we can solve for n_0

$$\hat{n}_0 = n \left(\frac{\text{Var}(S)}{\text{Var}(\theta_I)} - 1 \right), \quad (22)$$

where $\text{Var}(S)$ and $\text{Var}(\theta_I)$ are the variances of the summary statistic and the prior, respectively. We can use this equation to estimate n_0 in other prior-likelihood pairs. However, in these cases, we need to compute the summary statistic first because the prior and the likelihood may represent different quantities. For example, the prior in a beta-binomial is typically a probability between zero and 1, while the binomial likelihood is generally the distribution of the number of successes. Therefore, a summary statistic is needed to convert the marginalized likelihood of successes to the same *scale* as the prior probability. In the case of the beta-binomial, this can be easily achieved by dividing the number of successes k by the total sample n , such that $S = k/n$. However, calculating a summary statistic may not be always trivial. For such cases, we propose an indirect Markov chain Monte Carlo (MCMC) approach that does not require computing the summary statistic.

Computing n_0 indirectly via MCMC.—The indirect method involves computing n_0 for each parameter using an analytic method or MCMC. In most cases, n_0 can be estimated by generating new data from the prior and the likelihood function using equation (12), such that

$$\hat{n}_0 = n \left(\frac{\text{Var}(\theta_I)}{\text{Var}_X(\phi)} - 1 \right). \quad (23)$$

In Box 2, we summarize the steps involved in calculating n_0 using both the summary statistic and MCMC methods. (Appendix C provides the detailed steps for computing n_0 and its implementation in R.)

Case Study 1: Testing the Gaussian Approximation

In this section, we compare the distribution of the posterior mean estimated using the Gaussian approximation as defined in equation (16) to the posterior distribution estimated using traditional Bayesian updating through either conjugacy or MCMC. We tested the Gaussian approximation on 4 different examples: 1) beta-binomial, 2) gamma-exponential, 3) normal-Weibull, and 4) Dirichlet-multinomial prior-likelihood pairs. The distributions, their parameters, and the size of the data collection experiments are defined in Table 1.

The first experiment describes a discrete data collection exercise from a beta prior and a binomial data likelihood. The prior is conjugate to the data likelihood; therefore, the posterior is also beta distributed and can be computed analytically. The second experiment describes a gamma prior that is conjugate with the exponential data likelihood. The third experiment involves a normal prior that is not conjugate to the Weibull data likelihood. In this case, the posterior mean is computed via MCMC using JAGS.²⁶ The fourth example involves a Dirichlet prior with 3 components and a conjugate multinomial data likelihood.

In all cases, the prior sample size n_0 can be calculated directly using the parameters of the distributions through a closed-form solution except for the normal-Weibull case.²⁵ Therefore, we computed n_0 numerically using the indirect MCMC approach following the steps described in Box 2: 1) draw $\theta_I^{(1)}, \dots, \theta_I^{(m)}$ samples from the prior distribution of $\theta_I \sim \text{Normal}(\mu = 1, \sigma^2 = 0.04)$. 2) For each value of $\theta_I^{(i)}$, $i = \{1, \dots, m\}$, we took $j = \{1, \dots, n\}$ samples from the data likelihood, such that $x^{(ij)} \sim \text{Weibull}(v = 1, \lambda = 1/\theta^{(i)})$. We obtained q samples from the posterior distribution of $\theta|x^{(j)}$. Then, we calculated the posterior mean for each value of $\theta_I^{(i)}$, such that $\phi^{(i)} = \frac{1}{q} \sum_{k=1}^q \theta^{(k)}|x^{(i)}$. Finally, we estimated $\hat{n}_0 \approx 25$ using equation (23).

Case Study 2: Calculating EVSI in a Markov Model

We computed the EVSI of different parameters of an economic evaluation using a Markov model to compare the performance of the GA to the 2MCS approach. We chose a Markov model because Markov models are commonly used in economic evaluations and generally involve nonlinear relations among some of the model parameters and the net benefits.²⁷ (Details of the Markov model are described in Appendix D.) In summary, the model simulated a cohort with a hypothetical genetic disorder (syndrome X). Syndrome X is mostly asymptomatic but may cause a sudden flare-up that may necessitate hospitalization. There are 3 alternatives available for preventing progression to permanent disability: A, B, and C, where C is standard care. If treatment fails, permanent disability occurs and quality of life (QoL) decreases from 1.0 to 0.8. For ease of calculations, we assumed that patients face a constant annual mortality rate of 0.044. In addition, syndrome X is assumed to be associated with a 0.5% increase in absolute mortality. Furthermore, if disabled, mortality rate is assumed to increase by an additional 1%. Given that all transition probabilities are

time invariant, we used a fundamental matrix solution²⁸ of the Markov model to speed up the computation time.

Four parameters in the model are uncertain. The parameters representing the mean number of visits for interventions A and B have a linear relation with the net monetary benefit (NMB). However, the probability of failing A and B has a nonlinear relation with their respective interventions. This is because Markov models are typically linear in state payoffs but nonlinear in the state transition probabilities.

The n_0 is 10 for all parameters, which can be easily confirmed from the distribution characteristics of these parameters and their corresponding data likelihoods. For example, we assumed that the mean number of hospital visits for intervention A is distributed as gamma($a = 10$, $b = 0.1$). Because, the likelihood follows a Poisson distribution, this prior contains information equivalent to $n_0 = 1/b = 10$. In addition, the probability of failing A is distributed as beta($\alpha = 2$, $\beta = 8$) and the $n_0 = \alpha + \beta = 2 + 8 = 10$ because the data likelihood follows a binomial distribution.

We compared the estimated EVSI and EVPPI using 2 approaches: GA with splines as the linear metamodel (GA) and Bayesian updating with 2-level Monte Carlo simulations (2MCS). Each EVSI estimation with 2MCS using the fundamental matrix solution of the Markov model took about 6 hours on 16 parallel cores.

Results

Results of Case Study 1: Testing the Performance of the Gaussian Approximation

Figure 2 compares the empirical cumulative density function (ECDF) of the preposterior distribution for the 3 univariate experiments in Table 1 computed with traditional Bayesian updating and GA using equation (16). GA asymptotically approximates the preposterior distribution obtained with classic Bayesian updating. The step function for the preposterior distribution in the beta-binomial case through Bayesian updating is explained by the discrete nature of the binomial distribution. Numerically, the Kolmogorov-Smirnov (KS) statistic was close to zero for most experiments, indicating that the preposterior distributions computed with both Bayesian updating and GA were similar. In addition, the similarity between the preposterior distributions increased as n increased because with higher n , the Gaussian assumption becomes more appropriate.

Figure 3 compares the GA to Bayesian updating for the Dirichlet-multinomial experiment. This figure presents both the marginal and joint preposterior distributions for pairs of components of the Dirichlet distribution. The GA provides a good approximation of these distributions compared to the Bayesian updating. In addition, the correlation coefficients for pairs of components were also very similar.

Results of the Case Study 2: Calculating EVSI in a Markov Model

Base case analysis.—In the base case analysis, intervention B is the optimal strategy with the highest expected NMB at a willingness to pay (WTP) threshold of \$50,000/quality-adjusted life year (QALY) (Appendix D). However, intervention B is optimal in less than

50% of the PSA scenarios, indicating that based on current information, the decision maker faces a 54% chance of some opportunity loss.

VOI.—The EVPI was \$5480, which sets the upper limit on future research involving all the parameters in the decision model with current information. In addition, the EVPPI for the individual parameters and their corresponding standard errors (SEs) for GA and 2MCS are shown in Table 2. The SEs of the GA were calculated following a Monte Carlo approach²⁹ by evaluating equation (18) on n sets of coefficients $\alpha^{d(i)}$, where $i = 1, \dots, n$ is sampled from a multivariate Normal distribution with mean $\hat{\alpha}^d$ and covariance $\hat{\Sigma}^d$. $\hat{\alpha}^d$ and $\hat{\Sigma}^d$ are the estimated parameters and their corresponding covariance matrix obtained from the mgcv package. This approach has been previously used in the context of VOI by Strong and others.¹⁴ The SEs for the 2MCS were computed following Ades and others.¹² The estimated EVPPI using the GA closely resembles that obtained with the 2MCS approach but with higher SEs, which is the result of the additional errors from the metamodel approximation.

A single EVSI evaluation using the GA technique took 0.02 seconds on a single core while the 2MCS approach took 10 minutes on 16 parallel cores. Despite the efficiency of using the fundamental matrix solution, conjugate priors, and 16 parallel processing cores to compute EVSI using the 2MCS, it was still many folds slower compared to the GA technique.

Figure 4 compares the EVSI computed using GA and 2MCS for different sample sizes ($n = \{0, 5, 10, 50, 100, 1000, 10000\}$) and the associated CIs. When the sample size was greater than 10, the GA performed similar to the 2MCS for all parameters. Both approaches show that as the sample size increase, EVSI approaches EVPPI.

The GA approach generally produces smoother EVSI curves, especially for small n . This is mainly because the data likelihood used in the Bayesian updating samples integers from discrete distributions, causing larger variation in the posterior mean for smaller n .

Correlation.—In VOI, all evidence must be used to accurately define prior uncertainties. Such evidence often translates into correlated parameters. To illustrate the importance of correlation, we tested how VOI varies at different levels of correlations between different pairs of parameters of interest. The level of correlation was induced using a previously described algorithm.^{30,31} Briefly, this correlation induction involves taking a large number of samples from the independent prior distributions of the parameters and then sorting them in a way that induces the desired correlation structure. Because this algorithm only involves sorting the parameters, it preserves their marginal distributions. The algorithm can be extended to multiple correlated parameters.³² In case study 1, we illustrated the performance of the GA in approximating the joint preposterior distribution in the case of a Dirichlet prior and multinomial data likelihood. Furthermore, the GA can be easily extended to models that involve multiple correlated parameters that are not jointly sampled from the same multivariate distribution.

In Figure 5, we examine the effect of correlations (ρ) on the EVSI for all possible pairs of parameters. Correlations among parameters of interest have a drastic impact on the EVSI in all cases. The maximum effect can be seen between the mean number of hospital visits for

intervention A and the probability of failing intervention A. The extremes of ρ in this figure may not be realistic but illustrate the importance that correlation has on EVPPI and EVSI.

Discussion

We propose a general approach to compute EVSI that consists of 2 components: 1) a linear metamodel between θ_J and the opportunity loss L and 2) a Gaussian approximation of the posterior mean of the data collection experiments involving a set of parameters of interest θ_J . In a first case study, we illustrated the GA performance in terms of accuracy in 4 numerical exercises involving conjugate, nonconjugate, univariate, and multivariate priors. The second case study compared the accuracy and efficiency of the GA to the traditional Bayesian updating via the 2MCS approach in estimating EVPPI and EVSI of an economic evaluation using a Markov model with nonlinear relations between the net benefits and the non-Gaussian prior parameters.

The GA extends our previous efforts¹³ to calculate EVSI to a wide range of prior and likelihood combinations, multiple correlated parameters, and multiple strategies. The GA is an alternative to the traditional Bayesian updating process that can be extended to other non-Gaussian prior-likelihood pairs because it only requires the prior distribution. The GA does not impose any distributional form on the prior or the likelihood directly, and it works by generating a preposterior distribution that is the result of shrinking the prior toward its mean similar to updating a Gaussian prior using a Gaussian data likelihood. In addition, the GA generally offers additional computational gains over previous methods. For example, for each set of parameters, the analyst can use the same metamodel for various sample sizes n unlike previous approaches that may require refitting the metamodel for different n .¹⁴

Similar to Jalal and others,¹³ Strong and others,¹⁴ and Menzies,¹⁵ this study also adopts a regression metamodeling approach to facilitate the computation in EVSI. The linear regression metamodel approach is generally fast and simple to implement because it only requires the PSA data set and there is no need to rerun the simulation model after the PSA data set has been created, similar to previous approaches.^{13–15} In addition, when $n \rightarrow \infty$, our EVSI equation translates to an EVPPI equation that is similar to Strong and others³³ and Madan and others³⁴ using splines.

The GA approach simplifies the Bayesian updating process, which, in combination with the linear regression metamodeling, allows the posterior mean of the opportunity loss to be expressed as a function of the posterior mean values of the individual parameters. The GA approach can be used with various forms of regression metamodeling (e.g, simple linear or spline metamodels). In this article, we chose splines as a linear metamodel because splines are generally more flexible metamodels than simple linear metamodels and are easily implemented with mathematical and statistical software, such as R.^{14,15} The `mgcv` package contains algorithms to choose an appropriate number of basis functions. However, splines can fail if the number of basis functions in the spline exceeds the number of PSA samples, especially if multiple interactions of a large number of parameters are desired.¹⁴ In such cases, researchers can choose to explicitly define a simpler linear metamodel or a polynomial metamodel with fewer terms than the PSA samples. The choice of an

appropriate metamodeling technique may improve the accuracy of VOI computations, but researchers must always balance increased accuracy v. overfitting and should choose a metamodel that neither oversmooths nor overfits the data. This is an important step in all methods that use metamodels to compute either EVPPI^{13,33,34} or EVSI.^{13–15} Fortunately, many techniques (e.g., cross-validation) can be used to reduce model overfitting if overfitting is particularly important.

Since the prior is the source of all information in VOI, the GA assumes that all the randomness in ϕ is attributable to randomness in θ_I , resulting in ϕ and θ_I to be perfectly correlated. This assumption allowed us to express ϕ as a function of θ_I and extend the GA approach to cases in which the prior parameters are correlated as we illustrated above. This approach is different from traditional Bayesian updating because in traditional Bayesian updating, we assume that X is observed. Thus, the variability in the preposterior becomes a function of the variability in the prior and the data likelihood, which are often sampled from independent random distributions.

Theoretically, n_0 and n must be sufficiently large for the GA to perform well in approximating the preposterior distributions in non-Gaussian prior-likelihood cases. We tested the performance of this approximation numerically on various conjugate, nonconjugate, univariate, and multivariate non-Gaussian distributions when n and n_0 were as small as 10. In all cases, the GA approach performed well in approximating the Bayesian updating process. In addition, Pratt and others³⁵ have shown that when the sample size is *large enough* and its variance is known and finite, the Bayesian updating can be approximated by a Gaussian-Gaussian Bayesian update regardless of the nature of the functional form of the likelihood due to the central limit theorem.³⁵

However, if the sample sizes are particularly small or if the priors are highly skewed, the GA approach may not perform well. In these cases, researchers may conduct a full Bayesian updating or adopt other techniques such as transforming the parameters to be less skewed. For example, a log transformation can be used for rates or a logit transformation can be used for probabilities, but care must be given in all cases to compute n_0 accordingly.

A potential speed-limiting task in our approach is computing n_0 for complex study designs that require MCMC. Fortunately, n_0 needs to be computed only once for each parameter of interest, and then it can be reused as long as the prior distributions and the data likelihoods are not altered. As we have shown above, there are closed-form solutions for n_0 for some commonly used priors and data likelihood experiments. We illustrated the case of beta-binomial, gamma-exponential, gamma-Poisson, and Dirichlet-multinomial. We also proposed 2 novel approaches to calculate the value of n_0 from the prior and data likelihood using a GA through either a summary statistic or MCMC methods. In addition, numerical techniques have been proposed to calculate n_0 via Monte Carlo simulations.²⁵

In summary, the GA technique has several advantages: 1) it requires only the prior and the prior sample size n_0 , which can be estimated in various ways; 2) it can be used with different types of linear metamodels; 3) it is simple and efficient to compute compared to the

traditional Bayesian updating; and 4) it can be applied when the model parameters are correlated or interact.

Conclusion

EVSI is one of the most powerful concepts in medical decision making because it can inform future research, resource allocation, and data collection study designs. We propose a general GA approach to estimate EVSI that addresses some of the challenges associated with traditional EVSI calculations, especially when complex Bayesian updating of the prior uncertainties is required and when the prior evidence suggests that these uncertainties are correlated. The GA approach relies on defining the prior sample size n_0 , and it extends our prior efforts¹³ to multiple strategies and various distributional forms. We hope that this approach along with the other recent developments in computing EVSI can provide a useful set of tools for analysts to effectively conduct VOI.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank 2 anonymous reviewers for their valuable feedback that helped to improve this work substantially. We also thank Jay Bhattacharya, Nicky Welton, Torbjørn Wisløff, Yadira Peralta, and Rowan Iskandar for their helpful comments and suggestions on an earlier draft of this manuscript.

References

1. Claxton K and Posnett J. An economic approach to clinical trial design and research priority setting. *Health Econ.* 1996;5(6):513–24. [PubMed: 9003938]
2. Raiffa H. *Decision Analysis: Introductory Lectures on Choices under Uncertainty*. Reading, MA: Addison-Wesley; 1968.
3. Schlaifer R. *Probability and Statistics for Business Decisions*. New York: McGraw-Hill; 1959.
4. Schlaifer R and Raiffa H. *Applied Statistical Decision Theory*. 1961. Harvard Business School, Cambridge, MA
5. Brennan A and Kharroubi SA. Efficient computation of partial expected value of sample information using Bayesian approximation. *J Health Econ.* 2007;26(1):122–48. [PubMed: 16945438]
6. Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *J Health Econ.* 1999;18(3):341–64. [PubMed: 10537899]
7. Claxton K, Lacey LF and Walker SG. Selecting treatments: a decision theoretic approach. *J R Stat Soc.* 2000;163(2):211–25.
8. Claxton K, Cohen JT and Neumann PJ. When is evidence sufficient? *Health Affairs.* 2005;24(1):93–101. [PubMed: 15647219]
9. Steuten L, van de, Wetering G, Groothuis-Oudshoorn K and Retèl V. A systematic and critical review of the evolving methods and applications of value of information in academia and practice. *PharmacoEconomics.* 2013;31(1):25–48. [PubMed: 23329591]
10. Brennan A, Chilcott J, Kharroubi S and O'Hagan A. A two level Monte Carlo approach to calculation expected value of sample information: how to value a research design. Presented at the 24th Annual Meeting of the Society for Medical Decision Making; 10 20, 2002; Baltimore, MD.
11. Brennan A and Kharroubi SA. Expected value of sample information for Weibull survival data. *Health Econ.* 2007;1225:1205–25.

12. Ades A, Lu G and Claxton K. Expected value of sample information calculations in medical decision modeling. *Med Decis Making*. 2004;24(2):207–27. [PubMed: 15090106]
13. Jalal H, Goldhaber-Fiebert JD and Kuntz KM. Computing expected value of partial sample information from probabilistic sensitivity analysis using linear regression metamodeling. *Med Decis Making*. 2015;35(5):584–95. [PubMed: 25840900]
14. Strong M, Oakley JE, Brennan A and Breeze P. Estimating the expected value of sample information using the probabilistic sensitivity analysis sample: a fast, nonparametric regression-based method. *Med Decis Making*. 2015;35(5):570–83. [PubMed: 25810269]
15. Menzies NA. An efficient estimator for the expected value of sample information [published online April 24, 2015]. *Med Decis Making*.
16. Welton NJ and Thom HHZ. Value of information: we've got speed, what more do we need? *Med Decis Making*. 2015;35(5):564–6. [PubMed: 25840903]
17. Raiffa H and Schlaifer R. *Probability and Statistics for Business Decisions*. New York: McGraw-Hill; 1959.
18. Doubilet P, Begg CB, Weinstein MC, Braun P and McNeil BJ. Probabilistic sensitivity analysis using Monte Carlo simulation: a practical approach. *Med Decis Making*. 1985;5(2):157. [PubMed: 3831638]
19. Briggs A, Weinstein MC, Fenwick EL, Karnon J, Sculpher MJ and Paltiel D. Model parameter estimation and uncertainty analysis: a report of the ISPOR SMDM Modeling Good Research Practices Task Force Working Group-6. *Med Decis Making*. 2012;32(5):722–32. [PubMed: 22990087]
20. Kleijnen JPC. *Design and Analysis of Simulation Experiments*. 2nd ed. New York: Springer; 2015.
21. Jalal H, Dowd B, Sainfort F and Kuntz KM. Linear regression metamodeling as a tool to summarize and present simulation model results. *Med Decis Making*. 2013;33(7):880–90. [PubMed: 23811758]
22. Strong M, Oakley JE and Brennan A. Estimating multiparameter partial expected value of perfect information from a probabilistic sensitivity analysis sample: a nonparametric regression approach. *Med Decis Making*. 2014;34(3):311–26. [PubMed: 24246566]
23. Wood SN. *Generalized Additive Models: An Introduction*. Boca Raton, FL: Chapman and Hall/CRC Press; 2006.
24. Carlin BP and Louis TA. *Bayesian Methods for Data Analysis*. 3rd ed. Boca Raton, FL: CRC Press; 2009.
25. Morita S, Thall PF and Müller P. Determining the effective sample size of a parametric prior. *Biometrics*. 2008;64(2):595–602. [PubMed: 17764481]
26. Plummer M. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd International Workshop on Distributed Statistical Computing, 2003* 3 20 (Vol. 124, p. 125).
27. Sonnenberg FA and Beck JR. Markov models in medical decision making: a practical guide. *Med Decis Making*. 1993;13(4):322–38. [PubMed: 8246705]
28. Beck JR and Pauker SG. The Markov process in medical prognosis. *Med Decis Making*. 1983;3(4):419–58. [PubMed: 6668990]
29. Robert CP and Casella G. *Monte Carlo Statistical Methods*. 2nd ed. New York: Springer; 2004.
30. Iman RL and Conover WJ. A distribution-free approach to inducing rank correlation among input variables. *Commun Stat Simulation Computation*. 1982;11(3):311–34.
31. Jalal H and Kuntz KM. Building correlations among model parameters: a practical approach. Presented at the 35th Annual Meeting of the Society for Medical Decision Making; 10 19, 2013; Baltimore, MD.
32. Goldhaber-Fiebert JD and Jalal HJ. Some health states are better than others using health state rank order to improve probabilistic analyses [published September 16, 2015]. *Med Decis Making*.
33. Strong M, Oakley JE and Brennan A. Estimating multiparameter partial expected value of perfect information from a probabilistic sensitivity analysis sample: a nonparametric regression approach. *Med Decis Making*. 2014;34(3):311–26. [PubMed: 24246566]

34. Madan J, Ades AE, Price M, et al. . Strategies for efficient computation of the expected value of partial perfect information. *Med Decis Making*. 2014;34:327–42. [PubMed: 24449434]
35. Pratt JW, Raiffa H and Schlaifer R. *Introduction to Statistical Decision Theory*. Cambridge, MA: MIT Press; 1995.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Box 1

Summary of the Steps and R Code for Computing
EVSI Using Gaussian Approximation from a PSA
Data Set for a Parameter of Interest θ_I from a New

Sample n

1. Conduct probabilistic sensitivity analysis (PSA) and load the PSA data set.


```
library(mgcv); library(matrixStats)
psa <- read.csv("example_psa_dataset.csv")
n.sim <- nrow(psa) # number of simulations in psa
theta_I <- psa[, 1] # parameter of interest is column 1
nmb <- psa[, 5:7] # the strategies' net monetary
benefits are columns 5 through 7
```
2. Determine the optimal strategy d^* .


```
d.star <- which.max(colMeans(nmb))
```
3. Compute the opportunity loss $L(d, \theta)$.


```
loss <- nmb - nmb[, d.star]
```
4. Estimate a linear metamodel for the opportunity loss of each d strategy, L_d , by regressing them on the spline basis functions of θ_I .


```
lmm1 <- gam(loss[, 1] ~ s(theta_I))
lmm2 <- gam(loss[, 2] ~ s(theta_I))
lmm3 <- gam(loss[, 3] ~ s(theta_I))
```
5. Compute EVPPI using the estimated losses for each d strategy, \hat{L}_d , and applying equation (20).


```
Lhat <- cbind(lmm1$fitted, lmm2$fitted, lmm3$fitted) #
estimated losses
evppi <- mean(rowMaxs(Lhat)) # evppi equation
```
6. Load the predict.ga function.


```
source(GA_functions.R)
```
7. Compute the predicted loss for each d strategy, \hat{L}_d , given the prior sample size (n_0) and new sample size (n).


```
Ltilde1 <- predict.ga(lmm1, n = n, n0 = n0)
Ltilde2 <- predict.ga(lmm2, n = n, n0 = n0)
Ltilde3 <- predict.ga(lmm3, n = n, n0 = n0)
loss.predicted <- cbind(Ltilde1, Ltilde2, Ltilde3)
```
8. Compute EVSI using equation (18)


```
evsi <- mean(rowMaxs(loss.predicted)) # evsi equation
```

Appendix B provides example R codes for implementing this approach when there are multiple parameters and unbalanced data collection study designs.

Box 2

Summary of Steps for Computing n_0 from the Prior
with Gaussian Approximation Using Summary
Statistic and Markov Chain Monte Carlo (MCMC)
Approaches

A. Summary statistic

1. Take $i = 1, \dots, m$ samples from the prior distribution of θ_i .
2. Take n samples from the likelihood distribution of $x \mid \theta_i^{(i)}$.
3. Compute a summary statistic \mathcal{S} from the n samples of $x \mid \theta_i^{(i)}$.
4. Finally, estimate \hat{n}_0 using equation (22).

B. MCMC

1. Take $i = 1, \dots, m$ samples from the prior distribution of θ_i .
 2. Take n samples from the likelihood distribution of $x \mid \theta_i^{(i)}$.
 3. Obtain the posterior $\theta_i \mid x$ using a specialized package such as JAGS.
 4. Compute the posterior mean ϕ .
 5. Finally, estimate \hat{n}_0 using equation (23).
-

Appendix C provides examples in R for implementing these approaches.

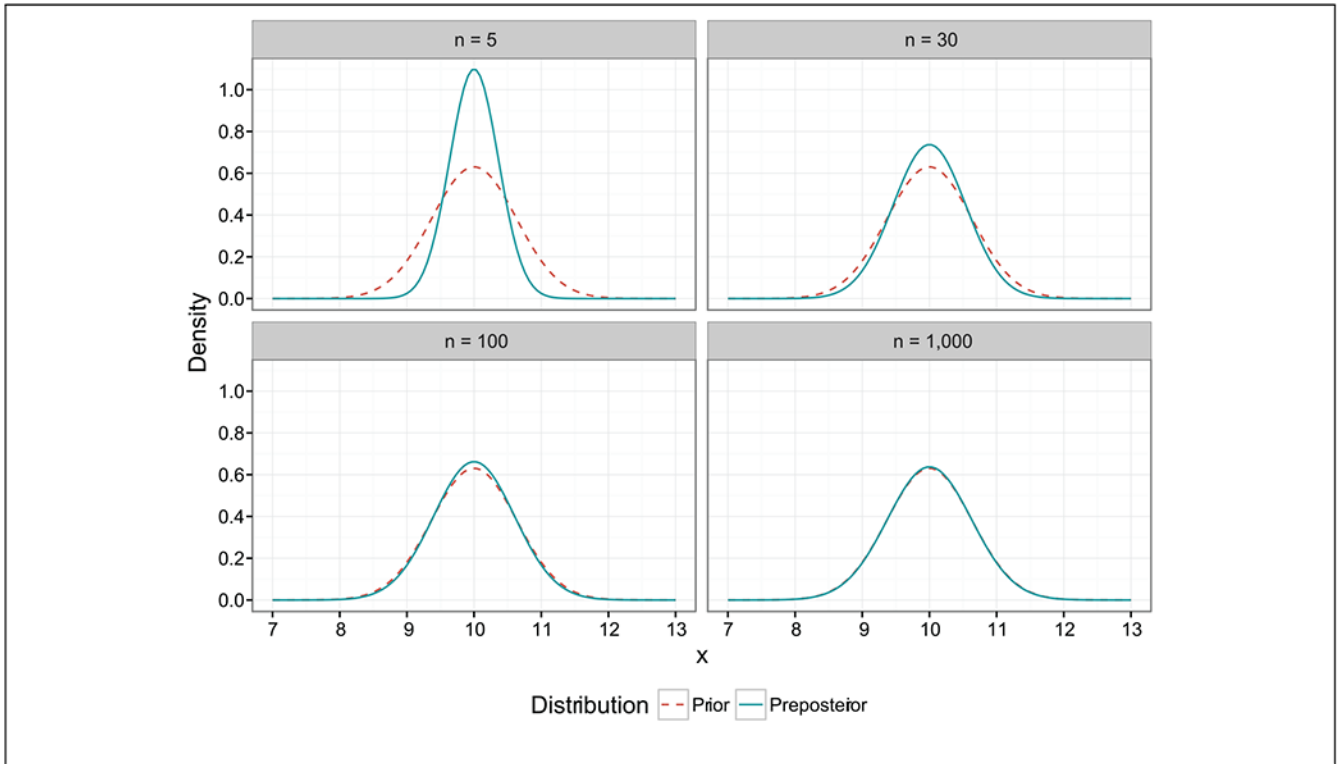


Figure 1. Gaussian preposterior and Gaussian prior distribution for different sample sizes n . The preposterior mean is the same for different n s. As n increases, the preposterior variance increases until it approximates that of the prior.

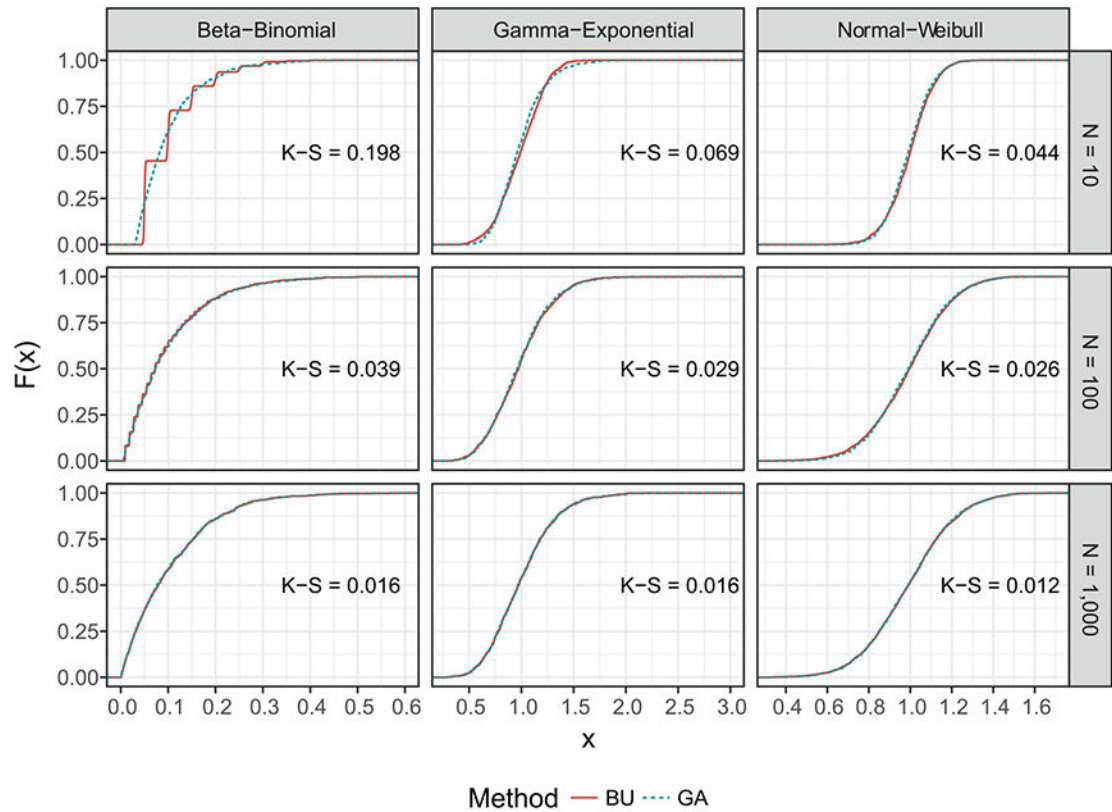


Figure 2. Comparison of the empirical cumulative distribution function (ECDF) of the Gaussian approximation (GA) to the traditional Bayesian updating (Bayesian updating) for 3 univariate data collection experiments: beta-binomial, gamma-exponential, and normal-Weibull prior-likelihood pairs. The posterior mean (ϕ) is shown on the x-axis and the ECDF ($F(x)$) at each value of ϕ is shown on the y-axis. The Kolmogorov-Smirnov (K-S) statistic quantifies a distance between the empirical distribution functions of 2 samples, where a smaller value provides stronger evidence that both samples come from the same distribution.

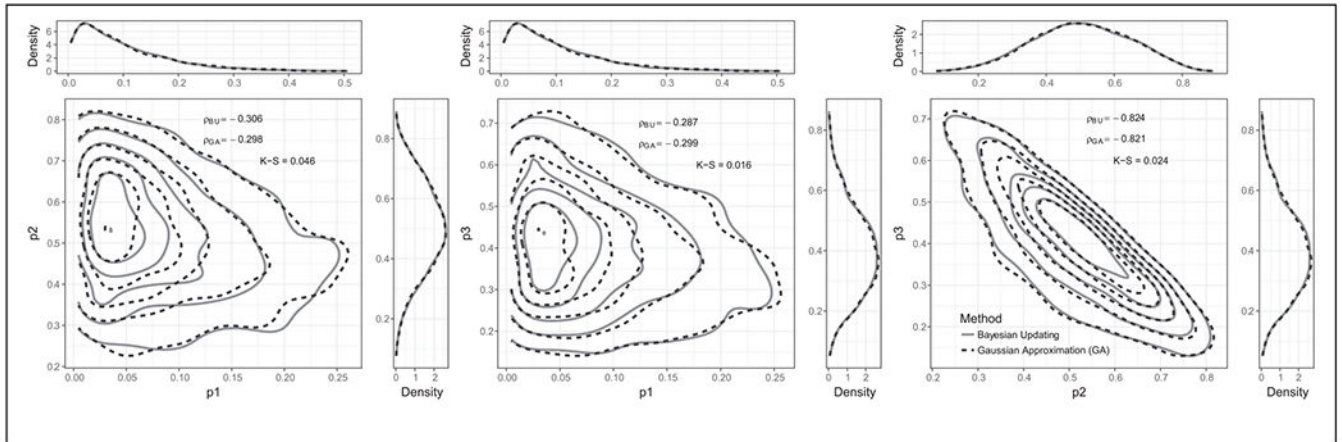


Figure 3.

Joint preposterior distribution for pairs of parameters in the fourth experiment involving a Dirichlet-multinomial prior-likelihood computed using the Gaussian approximation (GA) and traditional Bayesian updating. The marginal preposterior distributions are shown in the upper and right-hand side margins. ρ_{BU} and ρ_{GA} are the correlation coefficients between the prior component pairs using Bayesian updating and GA, respectively. The Kolmogorov-Smirnov (K-S) statistic quantifies a distance between the empirical distribution functions of 2 samples, where a smaller value provides stronger evidence that both samples come from the same distribution.

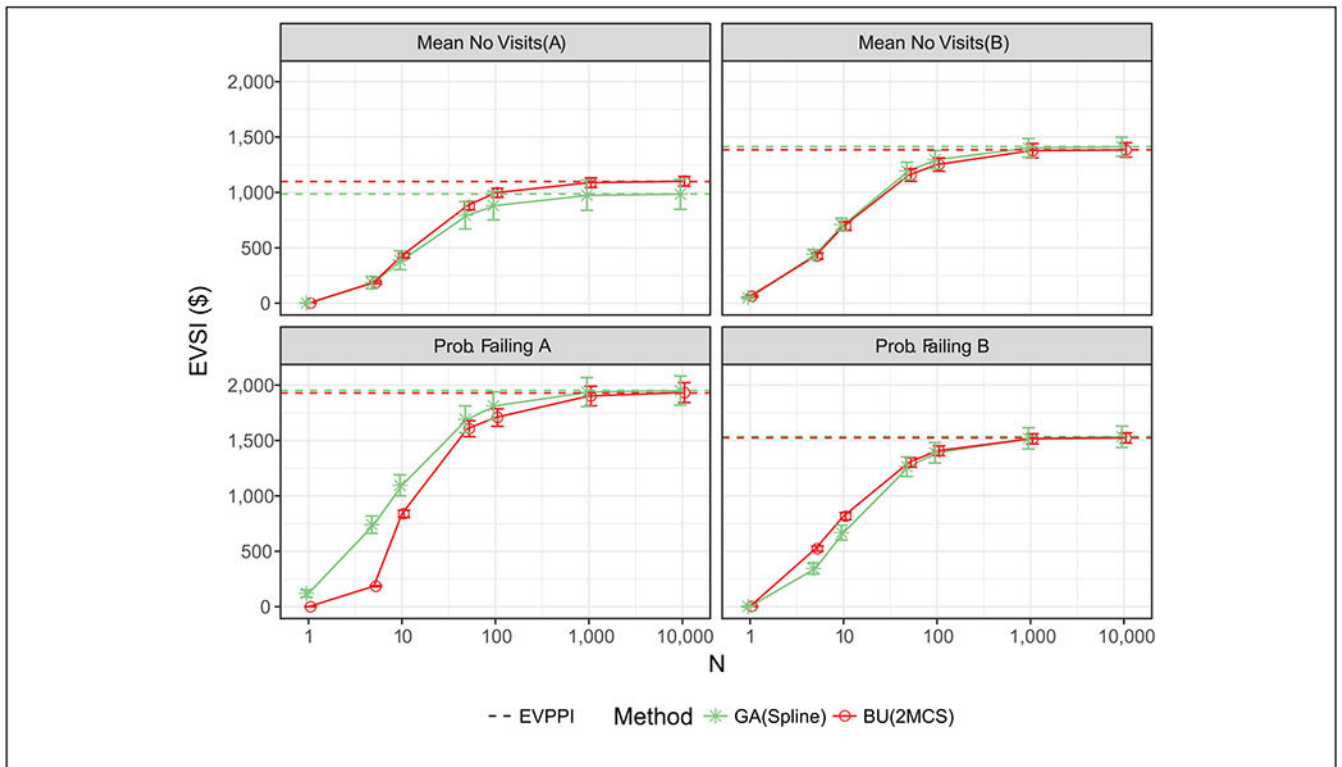


Figure 4. Expected value of sample information (EVSI) of each parameter for various sample sizes using Gaussian approximation (GA) with splines and traditional Bayesian updating with 2-level Monte Carlo simulations (2MCS) with their corresponding confidence intervals. The expected value of partial perfect information (EVPPI) using GA and Bayesian updating (BU) are also shown with the horizontal lines.

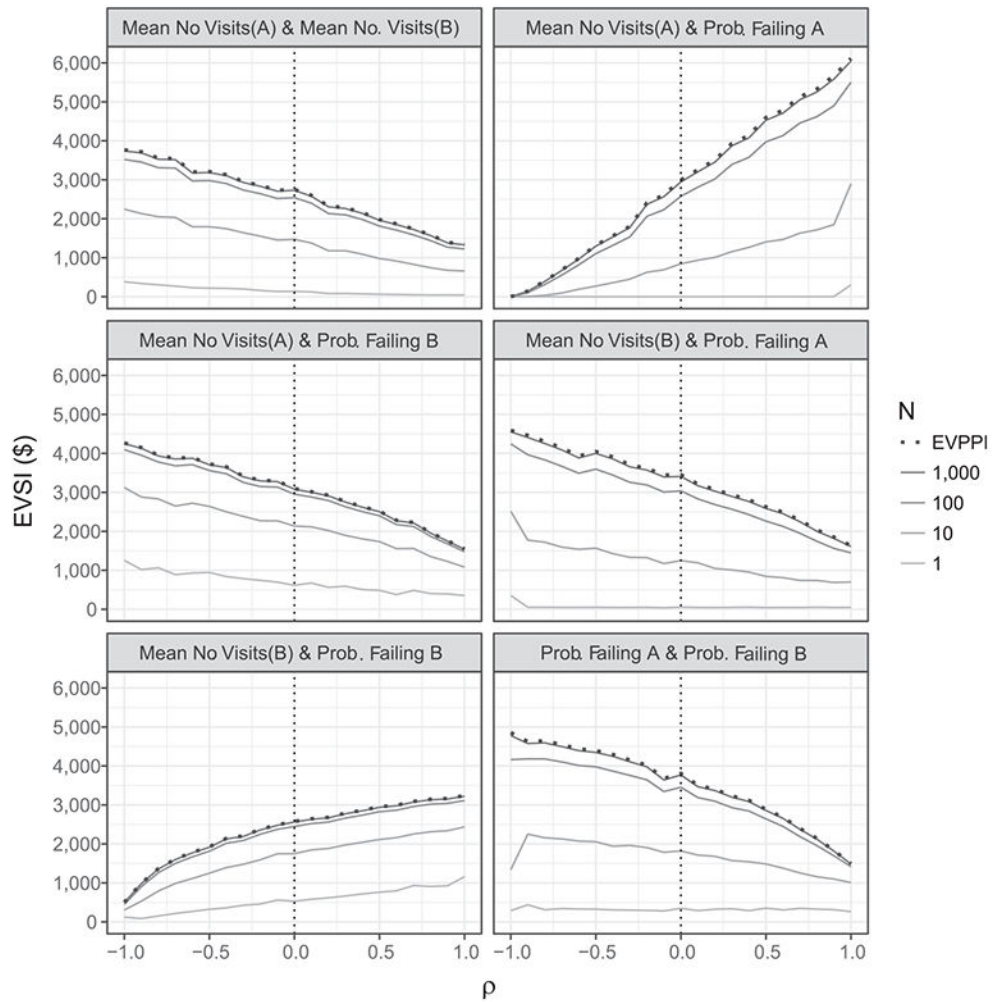


Figure 5. Impact of correlation on expected value of sample information (EVSI). This figure shows the change in EVSI for various sample sizes n for each pair of parameters in case study 2 over different levels of correlation between the parameters. The dotted vertical line represents the case where the parameters are independent. EV PPI, expected value of partial perfect information.

Traditional Bayesian Updating Experiments for Testing the Performance of the Gaussian Approximation

Table 1

	Experiment 1: Beta-Binomial	Experiment 2: Gamma-Exponential	Experiment 3: Normal-Weibull	Experiment 4: Dirichlet-Multinomial
Prior	$p \sim \text{beta}(a=1, \beta=9)$	$\lambda \sim \text{gamma}(a=10, b=0.1)$	$\theta \sim \text{normal}(\mu=1, \sigma^2=0.04)$	$p \sim \text{Dirichlet}(\alpha=[1, 5, 4])$
Likelihood	$x \sim \text{binomial}(n, p)$	$x \sim \text{exponential}(\lambda)$	$x \sim \text{Weibull}(v=1, \lambda=1/\theta)$	$x \sim \text{multinomial}(n, p)$
Posterior	$p x \sim \text{beta}(a+x, \beta+n-x)$	$\lambda x \sim \text{gamma}(a+n, b/(1+\sum_{i=1}^n x_i))$	Non-conjugate	$p x \sim \text{Dirichlet}(\alpha+x)$
\hat{n}_0 from the prior	$\hat{n}_0 = \alpha + \beta = 10$	$\hat{n}_0 = a = 10$	$\hat{n}_0 \approx 25$ (using MCMC)	$\hat{n}_0 = \sum \alpha_i = 10$
n	$n = \{10, 100, 1000\}$	$n = \{10, 100, 1000\}$	$n = \{10, 100, 1000\}$	$n = 100$
Univariate v. multivariate	Univariate	Univariate	Univariate	Multivariate

Estimated Expected Value of Partial Perfect Information for Each Parameter with Spline Regression and with 2-Level Monte Carlo Simulations (2MCS)

Table 2

Parameter	Spline Regression		2MCS	
	Mean (\$)	Standard Error (\$)	Mean (\$)	Standard Error (\$)
Mean number of visits (intervention A)	985	70	1098	22
Mean number of visits (intervention B)	1414	44	1384	33
Probability of failing A	1952	67	1928	46
Probability of failing B	1535	49	1524	24