Strategies for Efficient Computation of the Expected Value of Partial Perfect Information

Jason Madan, PhD, Anthony E. Ades, PhD, Malcolm Price, PhD, Kathryn Maitland, PhD, Julie Jemutai, BSc, Paul Revill, MSc, Nicky J. Welton, PhD

Expected value of information methods evaluate the potential health benefits that can be obtained from conducting new research to reduce uncertainty in the parameters of a cost-effectiveness analysis model, hence reducing decision uncertainty. Expected value of partial perfect information (EVPPI) provides an upper limit to the health gains that can be obtained from conducting a new study on a subset of parameters in the cost-effectiveness analysis and can therefore be used as a sensitivity analysis to identify parameters that most contribute to decision uncertainty and to help guide decisions around which types of study are of most value to prioritize for funding. A common general approach is to use nested Monte Carlo simulation to obtain an estimate of EVPPI. This approach is computationally intensive, can lead to significant sampling bias if an inadequate number of inner samples are obtained, and incorrect results can be obtained if correlations between parameters are not dealt with appropriately. In this article, we set out a range of methods for estimating EVPPI that avoid the need for nested simulation: reparameterization of the net benefit function, Taylor series approximations, and restricted cubic spline estimation of conditional expectations. For each method, we set out the generalized functional form that net benefit must take for the method to be valid. By specifying this functional form, our methods are able to focus on components of the model in which approximation is required, avoiding the complexities involved in developing statistical approximations for the model as a whole. Our methods also allow for any correlations that might exist between model parameters. We illustrate the methods using an example of fluid resuscitation in African children with severe malaria. ***Key words:*** *value-of-information; Bayesian methods; cost-effectiveness analysis.* ***(Med Decis Making 2014;34:327–342)***

Value of information (VoI) methods provide a coherent decision-theoretic approach to research prioritization.[1,2] When making a decision under imperfect information, it is possible that the expected optimal decision is wrong, and this possibility will be associated with an expected loss. VoI methods define the value of research in terms of the expected reduction in this expected loss resulting from the additional information gained. This is derived from the expected impact of the proposed study on parameter uncertainty. It is therefore natural to take a Bayesian perspective where prior beliefs about parameters are ''updated'' by incorporating new evidence to form posterior beliefs.[3] The expected loss given current information is also known

Received 31 August 2012 from the School of Social and Community Medicine, University of Bristol, Bristol, UK (JM, AEA, MP, NJW); Department of Medicine, Imperial College, London, UK (KM); KEMRI-Wellcome Trust Research Programme, Kilifi, Kenya (KM, JJ); Centre for Health Economics, University of York, York, UK (PR); Warwick Medical School, University of Warwick, Coventry, UK (JM); and School of Health and Population Sciences, University of Birmingham, Birmingham, UK (MP). JM and AEA were supported by the MRC ConDuCT Hub for Trials Methodology Research (grant G0800800), and NJW was supported by an MRC Methodology Research Fellowship (grant G0802413). Revision accepted for publication 25 September 2013.

© The Author(s) 2014
Reprints and permission:
http://www.sagepub.com/journalsPermissions.nav
DOI: 10.1177/0272989X13514774

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

Address correspondence to Jason Madan, PhD, Warwick Clinical Trials Unit, Division of Health Sciences, Warwick Medical School, The University of Warwick, Coventry CV4 7AL, UK; e-mail: j.j.madan@warwick.ac.uk.

as the expected value of perfect information (EVPI), because it is the amount the decision maker should be willing to pay to eliminate all parameter uncertainty in the decision.[1,2] Where there are multiple sources of uncertainty, it is possible to calculate the EVPI on a subset of focal parameters, termed the *expected value of partial perfect information*, or alternatively, *expected value of perfect parameter information* (EVPPI). EVPI and EVPPI provide an upper bound to the value of any proposed future study. Several authors have highlighted the use of these methods in health technology assessment in relation to sensitivity analysis and research prioritization, and we refer interested readers to this literature for a more complete discussion of the role of value-of-information analysis in health technology assessment.[4,5]

EVPPI involves an inner expectation nested within an outer expectation, which can be challenging to estimate; a common approach is nested Monte Carlo simulation.[6] There are 3 main issues with the nested simulation approach. First, obtaining estimates of adequate precision can be computationally expensive.[7] Second, estimates will have significant upward bias if insufficient simulation sample sizes are used for the inner loop.[8] Third, it can be difficult to generate unbiased EVPPI estimates when correlations exist between parameters, unless the joint distribution of the correlated parameters takes a known parametric form.[7] Conditions for the net benefit function have been identified that, if satisfied, allow EVPPI to be calculated in a single step by replacing the nonfocal parameters with their expected means.[9] In this article, we extend the work of Ades and others,[9] describing methods that can be used to avoid nested simulation in a broad range of situations, through careful consideration of the structure of the net benefit function and the relationships between its parameters. We illustrate these methods using a case study involving fluid resuscitation in African children with severe malaria, followed by a discussion of the advantages of the methods compared with other approaches and limitations resulting from the assumptions made.

Motivating Example

Malaria accounts for up to 1 million deaths annually in children younger than 5 years living in Sub-Saharan Africa. Children hospitalized with cerebral malaria (coma) have increased case fatality (18%–21%),[10] and survivors are at heightened risk of long-term neurological sequelae (NS).[11,12] There are clinical arguments suggesting that fluid resuscitation could reduce mortality in African children.[13] Furthermore, pilot studies, involving 3 fluids (saline, albumin, and gelofusine), provided limited but promising support for this hypothesis.[13–15] Following these studies, a large multinational trial, the FEAST trial, was initiated to provide definitive evidence on the efficacy of fluid resuscitation.[16]

Table 1 Assumptions Involved in Constructing the Net Benefit Function

- A1: Fatalities from severe malaria occur within the first few days following hospitalization.
- A2: Mortality differences between fluids are assumed additive on the log-odds scale.
- A3: Survivors and nonsurvivors receive identical treatment, except that survivors have a longer mean stay in the hospital. This additional stay is associated with an incremental cost of C^H.
- A4: Neurological sequelae (NS) in survivors become apparent within 28 d of discharge.
- A5: Giving fluid resuscitation to children who would survive anyway has minimal impact on their likelihood of developing NS.
- A6: Improved survival from treatment is associated with a change d^S in the log-odds of NS among those who are ''saved'' by treatment (see main text for further explanation).
- A7: A proportion p^L of NS cases at 28 d will persist long term.
- A8: Short-term cases of NS have no impact on treatment costs or quality of life.
- A9: Survivors who are free of long-term NS have a mean quality-adjusted life expectancy of q^M.
- A10: Long-term NS is associated with a reduction in quality-adjusted life expectancy of Q^S and increased discounted lifetime management costs of C^S. These values are assumed to be the same irrespective of fluids received.
- A11: Parameters are assumed to be independent unless correlations are induced between them through joint estimation from a common data source.

We developed an economic model to assess the cost-effectiveness of fluid resuscitation strategies, taking a health services perspective in which the relevant costs and benefits are those relating to the individual patient (irrespective of who actually provides the funds for health care). The model is based on the assumptions given in Table 1, which were informed by consultation with clinical experts advising the FEAST trial. Assumption A6, used to derive the probability of NS in those receiving fluids, requires further explanation. If fluid resuscitation reduces mortality, there will be a subgroup of survivors who would have died without fluids. Assumption A6 allows for the possibility that this subgroup is more severely affected by malaria than those who would

have survived without fluids and therefore are at greater risk of NS. Note that we do not assume this a priori; the model allows for the possibility that fluids increase mortality or that improved survival is associated with reduced NS risk.

There are 4 possible outcomes for patients in the model: death, NS-free survival, survival with short-term NS, and survival with long-term NS. Table 2 lists the probability of each outcome with or without fluid resuscitation, and a full list of model parameters is given in Table 3. To assess the cost-effectiveness of treatment j, we calculate its incremental net benefit $B_j(\theta)$ relative to no fluid, that is, the additional costs incurred and health utility gained above those observed without fluid resuscitation. This includes the cost C_j^F of fluid j and additional costs and benefits related to changes in the proportion achieving each of the outcomes in Table 2, that is,

$$B_j(\theta) = \left(\frac{\exp(\alpha)}{(1+\exp(\alpha))} - \frac{\exp(\alpha + d_j^M)}{(1+\exp(\alpha + d_j^M))}\right) (Wq^M - C^H - p^L\left(\frac{p^B \exp(d^S)}{(1+p^B \exp(d^S))}\right)(Wq^S + C^S)) - C_j^F \quad (1)$$

where $\theta = \left\{\alpha, d_j^M, q^M, q^S, p^L, d^S\right\}$, W = willingness to pay per quality-adjusted life-year.

This model (eq. 1) was designed to estimate the cost-effectiveness of fluids once FEAST trial results were available. However, it can also be used in an EVPPI analysis to provide insight on the potential value of the FEAST trial and other studies that could be undertaken to inform parameters in eq. 1. This requires quantifying the uncertainty around the parameters implied by information available at the time the FEAST trial was designed (Table 3). Values for C^H, C^F, and q^M were based on health economic analysis carried out by several of the authors alongside the FEAST trial.[16] Data on NS-related economic parameters are sparse; our chosen values are toward the upper end of what were considered plausible values, for reasons that are explored in the Discussion section.

Table 3 also lists the type of future studies that can provide information on each of the uncertain model parameters. We assumed that baseline mortality in a randomized controlled trial (RCT) population would be different from that in the decision populations, so that a different study type would be needed to provide information on baseline and relative effect parameters. In practice, if the 2 populations are considered sufficiently similar, all of these parameters

Table 2 Possible Outcomes for Patients in the Economic Model, with the Probability of Each Outcome and Associated Health Utility and Costs

Patient Outcome	Health Utility (QALYs)	Cost	Probability without Fluids	Probability with Fluid j
O1: Dies	0	C_j^F	$\frac{\exp(\alpha)}{(1+\exp(\alpha))}$	$\frac{\exp(\alpha + d_j^M)}{(1+\exp(\alpha + d_j^M))}$
O2: Survives, no NS	q^M	$C^H + C_j^F$	$\left(1 - \frac{\exp(\alpha)}{(1+\exp(\alpha))}\right)(1-p^B)$	$1 - \left(\frac{\exp(\alpha)}{(1+\exp(\alpha))}\right)(1-p^B) + \left(\frac{\exp(\alpha)}{(1+\exp(\alpha))} - \frac{\exp(\alpha + d_j^M)}{(1+\exp(\alpha + d_j^M))}\right)\left(1 - \frac{p^B \exp(d^S)}{(1+p^B \exp(d^S))}\right)$
O3: Survives, NS at 28 d but not long-term	q^M	$C^H + C_j^F$	$\left(1 - \frac{\exp(\alpha)}{(1+\exp(\alpha))}\right)(1-p^L)p^B$	$1 - \left(\frac{\exp(\alpha)}{(1+\exp(\alpha))}\right)(1-p^L)p^B + \left(\frac{\exp(\alpha)}{(1+\exp(\alpha))} - \frac{\exp(\alpha + d_j^M)}{(1+\exp(\alpha + d_j^M))}\right)\left(\frac{p^B \exp(d^S)}{(1+p^B \exp(d^S))}\right)(1-p^L)$
O4: Survives with long-term NS	$q^M - q^S$	$C^H + C^S + C_j^F$	$\left(1 - \frac{\exp(\alpha)}{(1+\exp(\alpha))}\right)p^L p^B$	$1 - \left(\frac{\exp(\alpha)}{(1+\exp(\alpha))}\right)p^L p^B + \left(\frac{\exp(\alpha)}{(1+\exp(\alpha))} - \frac{\exp(\alpha + d_j^M)}{(1+\exp(\alpha + d_j^M))}\right)\left(\frac{p^B \exp(d^S)}{(1+p^B \exp(d^S))}\right)p^L$

Note: NS = neurological sequelae; QALY = quality-adjusted life-year.

Table 3 Values/Distributions Used for Parameters in the Economic Model and Study Designs That Would Provide Further Information on Them

Parameter	Description	Value/Distribution	Study Design to Provide Further Information
d_2^M	Effect of albumin on mortality	Posterior distribution generated by Bayesian evidence synthesis model	RCT including albumin and control arm
d_3^M	Effect of saline on mortality	Posterior distribution generated by Bayesian evidence synthesis model	RCT including saline and control arm
d_4^M	Effect of gelofusine on mortality	Posterior distribution generated by Bayesian evidence synthesis model	RCT including gelofusine and control arm
d^S	Change to NS risk in ''saved'' patients	Posterior distribution generated by Bayesian evidence synthesis model	RCT including fluid resuscitation arm(s) and control arm
C^F	Fluid cost (per patient)	\$1 (saline), \$35 (albumin), \$12.50 (gelofusine)	
C^H	Additional in-patient costs associated with survival	\$60 (based on 5 d at \$12/d)	
α	Log-odds of death without fluid resuscitation	Normal with implied median probability of death 25%, 95% CI 15%–40%	Cohort study with short-term follow-up
p^B	Probability of NS without fluid resuscitation	Beta (1,9) (mean 10%, 95% CI 0.6%–28.5%)	Cohort study with 28-d follow-up
p^L	Probability that NS will still be present at 6 mo conditional on NS observed at 28 d	Beta (1,1)	Cohort study with 6-mo follow-up
C^S	Long-term discounted costs of NS	\$20 000	
q^M	QALY loss per fatality	Normal with CHAR1 = 20 and s = 5	Cohort study on survivors without NS: long-term follow-up
q^S	QALY loss per case of NS	Truncated normal with CHAR1 = 5, s = 3.16, lower limit = 0	Cohort study on those with NS: long-term follow-up

Note: CI = confidence interval; NS = neurological sequelae; QALY = quality-adjusted life-year; RCT = randomized controlled trial.

could be estimated from an RCT. Below, we present generic methods that can be used to estimate EVPPI in a wide range of circumstances. We then illustrate their application for the fluid resuscitation example by considering the different future studies that we could run (Table 3) and the subset of parameters those studies would directly inform.

METHODS

Let $B_j(\theta)$be a general form for the incremental net benefit of intervention j given model parameters θ and j^* be the optimal decision given current information. We define φ as a subset of (focal) parameters on which we are considering collecting further information on in a new study, and φ^C as the set of remaining (nonfocal) model parameters. $EVPPI(\varphi)$ equals the expected gain in net benefit, given current information, from switching to the optimum treatment once φ is known with certainty:

$$EVPPI(\varphi) = \underset{\varphi}{E}\left[\max_j\left\{\underset{\varphi^C|\varphi}{E}\left[B_j(\theta) - B_{j*}(\theta)\right]\right\}\right]. \quad (2)$$

We describe 2 types of methods that can be used (if necessary in combination) to avoid the inner-simulation step when calculating $EVPPI(\varphi)$, which all involve finding a solution (exact or approximate) for the inner expectation $\underset{\varphi^C|\varphi}{E}\left[B_j(\theta) - B_{j*}(\theta)\right]$.

Methods in Which Expectations for (Functions of) φ^C Can Be Plugged in Directly

Method 1: Direct Substitution of Means of the Nonfocal Parameters in a Linear Net Benefit Function

This method is applicable when the net benefit takes the form

$$B_j(\varphi^C, \varphi) = \sum_i f_{i,j}(\varphi)\varphi_i^C,$$

where φ_i^C,the ith component of φ^C, is independent of any of the elements of φ in its linear coefficient $f_{i,j}(\varphi)$. Given this independence, it has been shown[4,9] that

$$\underset{\varphi^C|\varphi}{E}\left[B_j(\varphi^C, \varphi)\right] = \sum_i f_{i,j}(\varphi)E\left[\varphi_i^C\right]. \quad (3)$$

$EVPPI(\varphi)$can now be estimated in a single Monte Carlo simulation by substituting equation 3 into equation 2.

Method 2: Direct Substitution of Means of the Nonfocal Parameters in a Multilinear Net Benefit Function

This method is applicable when the net benefit can be expressed in the form

$$B_j(\varphi^C, \varphi) = \sum_i f_{i,j}(\varphi)g_{i,j}(\varphi^C),$$

where

- each function $g_{i,j}(\varphi^C)$ is a product of mutually independent elements of φ^C and
- each of the elements of $g_{i,j}(\varphi^C)$ is independent of the elements of φ in its linear coefficient $f_{i,j}(j, \varphi)$.

If these conditions are satisfied, it has been shown[4,9] that the expected value of each function $g_{i,j}(\varphi^C)$will equal the product of the expectations of its components, so that

$$\underset{\varphi^C|\varphi}{E}\left[B_j(\varphi^C, \varphi)\right] = \sum_i f_{i,j}(\varphi)g_{i,j}\left(E\left(\varphi_1^C\right), E\left(\varphi_2^C\right), \ldots\right). \quad (4)$$

$EVPPI(\varphi)$can now be estimated in a single Monte Carlo simulation by substituting equation 4 into equation 2.

Method 3: Reparameterization to Linearize the Net Benefit Function

This method is applicable when the net benefit can be expressed in the form

$$B_j(\varphi^C, \varphi) = \sum_i f_{i,j}(\varphi)\beta_{i,j}(\varphi^C)$$

where

- $\beta_{i,j}(\varphi^C)$ are any functions of the φ^C parameters and
- the elements of φ^C in each $\beta_{i,j}(\varphi^C)$are all independent of any of the elements of φ in its linear coefficient $f_{i,j}(\varphi)$.

Method 2 is not applicable if any of the $\beta_{i,j}(\varphi^C)$are not products of their components and/or include correlated nonfocal parameters . However, the second condition ensures the independence of each $\beta_{i,j}(\varphi^C)$from its linear coefficient $f_{i,j}(\varphi)$, which implies that

$$\underset{\varphi^C|\varphi}{E}\left[B_j(\varphi^C, \varphi)\right] = f_{i,j}(\varphi)\underset{\varphi^C}{E}\left[\beta_{i,j}(\varphi^C)\right]. \quad (5)$$

$EVPPI(\varphi)$can now be estimated in a single Monte Carlo simulation by substituting equation 5 into equation 2. This method has been used by Welton and others.[17]

Methods Involving Functions Approximating the Conditional Expectations of φ^C

Method 4: Taylor Series Approximations to the Net-Benefit Function

This method is applicable when net benefit takes the form

$$B_j(\varphi^C, \varphi) = \sum_i f_{i,j}(\varphi)\beta_{i,j}(\varphi_{-i}^C)h_{i,j}(\varphi_i^C, \varphi),$$

where

- $f_{i,j}(\varphi)$are arbitrary functions of the focal parameters and $h_{i,j}(\varphi_i^C, \varphi)$are arbitrary smooth nonlinear functions of φ and a single element φ_i^C of φ^C,
- $\beta_{i,j}(\varphi_{-i}^C)$are arbitrary functions of the remaining φ^C parameters (excluding φ_i^C), and
- φ_i^C, the elements of φ in $h_{i,j}(\varphi_i^C, \varphi)$, and the elements of φ^Cin $\beta_{i,j}(\varphi_{-i}^C)$ are mutually independent.

Because φ_i^C, φ_{-i}^C and the elements of φ *in* $h_{i,j}(\varphi_i^C, \varphi)$ are mutually independent, we can decompose the expectation:

$$\underset{\varphi^C|\varphi}{E}\left[\sum_i f_{i,j}(\varphi)\beta_{i,j}(\varphi_{-i}^C)h_{i,j}(\varphi_i^C, \varphi)\right] = \sum_i f_{i,j}(\varphi)\,E\left[\beta_{i,j}(\varphi_{-i}^C)\right]E\left[h_{i,j}(\varphi_i^C, \varphi)\right].$$

The expectation of each $\beta_{i,j}(\varphi_{-i}^C)$, if required, can be found using method 3. Because $h_{i,j}(\varphi_i^C, \varphi)$is nonlinear, its expectation conditional on φ will not equal $h_{i,j}\left(E\left(\varphi_i^C\right), \varphi\right)$, and we cannot use methods 1 to 3. Our strategy is instead to construct a function of $E\left(\varphi_i^C\right)$ that approximates the required conditional expectation and so can be used to replace the inner simulation step. One approach is to use Taylor series expansions.[9] The nth-order Taylor series expansion of $h_{i,j}(\varphi_i^C, \varphi)$ in the neighborhood of $E\left(\varphi_i^C\right)$ is given by

$$h_{i,j}(\varphi_i^C, \varphi) \simeq h_{i,j}(\boldsymbol{E}[\varphi_i^C], \varphi) + \sum_{r=1}^{n} \frac{1}{r!} \frac{d^r h_{i,j}(\boldsymbol{E}[\varphi_i^C], \varphi)}{d(\varphi_i^C)^r} (\varphi_i^C - \boldsymbol{E}[\varphi_i^C])^r.$$

As long as φ_i^C is independent of the elements of φ in $h_{i,j}(\varphi_i^C, \varphi)$, taking conditional expectations on both sides gives

$$\underset{\varphi_i^C|\varphi}{\boldsymbol{E}} \left[h_{i,j}(\varphi_i^C, \varphi)\right] \simeq h_{i,j}(\boldsymbol{E}[\varphi_i^C], \varphi) + \frac{1}{2} \frac{d^2 h_{i,j}(\boldsymbol{E}[\varphi_i^C], \varphi)}{d(\varphi_i^C)^2} \cdot Var[\varphi_i^C] + \ldots.$$

This can be used to give an approximation to the conditional expectation of the net benefit function; for example, a second-order approximation is

$$\underset{\varphi^C|\varphi}{\boldsymbol{E}} \left[B_j(\varphi^C, \varphi)\right] \approx \sum_i f_{i,j}(\varphi) \bar{h}_{i,j}^{T2}(\boldsymbol{E}[\varphi_i^C], \varphi) \boldsymbol{E}[\beta_{i,j}(\varphi_{-i}^C)], \quad (6)$$

where

$$\bar{h}_{i,j}^{T2}(\boldsymbol{E}[\varphi_i^C], \varphi) = h_{i,j}(\boldsymbol{E}[\varphi_i^C], \varphi) + \frac{1}{2} \frac{d^2 h_{i,j}(\boldsymbol{E}[\varphi_i^C], \varphi)}{d(\varphi_i^C)^2} \cdot Var[\varphi_i^C].$$

$EVPPI(\varphi)$can now be estimated in a single Monte Carlo simulation by substituting equation 6 into equation 2.

The form of the approximation will depend on the functional form of $h_{i,j}(\varphi_i^C, \varphi)$. For example, net benefit functions commonly consist of an absolute probability that is informed by parameters on the log-odds scale (baseline log-odds plus a log-odds ratio for treatment j), so that $h_{i,j}(\varphi_i^C, \varphi)$ is the inverse-logit function of the form

$$h_{i,j}({\varphi_i}^C, \varphi) = \frac{e^{({\varphi_i}^C + \varphi)}}{1 + e^{({\varphi_i}^C + \varphi)}}. \quad (7)$$

The second-order Taylor series approximation for the inverse-logit function is (see supplementary appendix):

$$\underset{\varphi^C|\varphi}{\boldsymbol{E}} \left[h(j, \varphi_i^C, \varphi)\right] \simeq H + \frac{H(1-H)(1-2H)Var[\varphi_i^C]}{2}$$

$$\text{where} \quad H = \frac{e^{(E[\varphi_i^C] + \varphi)}}{1 + e^{(E[\varphi_i^C] + \varphi)}}.$$

Occasionally, the absolute probability is informed by parameters on the cumulative log-log (clog-log) scale, so that function h is the inverse-clog-log function of the form

$$h_{i,j}({\varphi_i}^C, \varphi) = 1 - e^{-e^{({\varphi_i}^C + \varphi)}}.$$

In this case, it can be shown that the second-order Taylor approximation gives (see supplementary appendix)

$$\underset{\varphi^C|\varphi}{\boldsymbol{E}} \left[h(j, \varphi_i^C, \varphi)\right] \simeq 1 - H + \frac{H\ln(H)(\ln(H) - 1)Var[\varphi_i^C]}{2}$$

$$\text{where} \quad H = e^{-e^{(E[\varphi_i^C] + \varphi)}}.$$

When using Taylor series approximations, it is important to assess the accuracy of the approximation and to consider adaptations that improve accuracy, for example, the addition of higher-order terms. We have developed the following adaptation for this purpose. We split the prior distribution for φ_i^C into quantiles, Q, and find the mean, μ_Q, of φ_i^C conditional on lying in quantile Q. We then construct separate Taylor series approximations to $h_{i,j}({\varphi_i}^C, \varphi)$ within each quantile around the quantile-specific mean, μ_Q. It can be shown that the expected value of the average over this set of Taylor series approximations converges to the required quantity, $\underset{\varphi^C|\varphi}{\boldsymbol{E}} \left[h(j, \varphi_i^C, \varphi)\right]$ (details and proof in the supplementary appendix). Figure 1 illustrates the accuracy of Taylor series expansions in the case in which ${\varphi_i}^C$ is normally distributed, $h_{i,j}({\varphi_i}^C, \varphi)$ takes the form given in equation 6, and $E[\varphi + \varphi_i^C] = 2$. The accuracy of approximations based on Taylor series expansions around $E[\varphi_i^C]$ reduces as the variance of φ_i^C increases, and the fourth-order approximation is noticeably more accurate than the second-order approximation. However, averaging over second-order expansions around the interquartile means is more effective in improving accuracy than increasing the order of the single approximation. The Taylor series gives local approximations whose accuracy diminishes as the distance from the expansion point increases. Therefore, averaging over multiple expansions around a range of values reduces error in the approximation by restricting the range over which each approximation is made, and this is our recommended approach.

Method 5: Spline-Approximation Methods When There Are Correlations between Parameters This method is applicable when the net benefit takes the form

$$B_j(\varphi^C, \varphi) = \sum_i f_{i,j}(\varphi) h_{i,j}(\varphi^C, \varphi_i),$$

where

- $f_{i,j}(\varphi)$ are arbitrary functions of the focal parameters,
- $h_{i,j}(\varphi^C, \varphi_i)$ are arbitrary smooth nonlinear functions of φ^C and a single element φ_i of φ, and
- for at least one value of i, φ_i is correlated with one or more of the elements of φ^C in $h_{i,j}(\varphi^C, \varphi_i)$.

Here we can no longer use Taylor series expansions to estimate the conditional expectation of $h_{i,j}(\varphi^C, \varphi_i)$ in terms of the unconditional expectation of φ^C, because elements of the latter are correlated withφ_i. Instead, we aim to construct functions $\xi_{i,j}(\varphi_i)$ whose values are close approximations to $\underset{\varphi^C|\varphi}{E}\left[h_{i,j}(\varphi_i^C, \varphi)\right]$, so that

$$\underset{\varphi^C|\varphi}{E}\left[B_j(\varphi^C, \varphi)\right] \approx \sum_i f_{i,j}(\varphi)\xi_{i,j}(\varphi_i). \quad (8)$$

$EVPPI(\varphi)$ can then be estimated in a single Monte Carlo simulation by substituting equation 8 into equation 2.

An obvious choice for $\xi_{i,j}(\varphi_i)$ is a spline.[18] There are several different types of splines; we describe here the use of restricted cubic splines for this purpose. The restricted cubic spline $\xi_{i,j}(\varphi_i)$ takes the form

$$\xi_{i,j}(\varphi_i) = \gamma_{j,0} + \gamma_{j,1} + \sum_{n=1}^{N_\tau} \gamma_{j,n+1}(\max\,(0, \varphi_i - \tau_n))^3, \quad (9)$$

where $\{\tau_1, \ldots \tau_{N_\tau}\}$ are the N_τ knots that define the breakpoints of $\xi_{i,j}(\varphi_i)$. The parameters γ are chosen so that $\xi_{i,j}(\varphi_i)$ is linear for values of φ_i less than τ_1 or greater than $\tau_{N\tau}$. This requires a minimum of 3 knots. Furthermore, $\xi_{i,j}(\varphi_i)$ must be continuously differentiable, which places further restrictions on the γ. To construct these splines, we first require estimates of $\underset{\varphi^C|\varphi}{E}\left[h_{i,j}(\varphi^C, \varphi_i)\right]$ for a large set of values for φ_i spanning a plausible range. Once knots $\{\tau_1, \ldots \tau_{N_\tau}\}$ have been chosen, regression analysis can be used to derive coefficients γ defining the splines that fit the estimated conditional expectations most closely. This approach can be implemented by the following algorithm.

Step 1: Generate N_θ samples from the joint distribution of θ and estimate $h_{i,j}(\varphi^C, \varphi)$ for each sample and each intervention j. These should be the same set of samples that will subsequently be used to estimate $EVPPI(\varphi)$.

Step 2: Divide the range spanning the values of φ_i into M^B bins $\Phi_m = \{x_{m-1}, x_m\}, 2 \leq m \leq M^B$. We chose values for x_m so that there were an equal number of samples in each bin (this is not the only possible approach, e.g., bins could be of equal width).

Step 3: For each bin Φ_m :

a. Identify the subset of samples within θ for which the sampled value of φ_i lies within Φ_m.
b. Calculate the mean $\overline{\varphi}_{i,m}$ of all the sampled values of φ_i in the subset.
c. Calculate the mean $\overline{h}_{i,j,m}$ of all the sampled values of $h_{i,j}(\varphi^C, \varphi_i)$ in the subset.

Step 4: Choose a vector of knots $\{\tau_1, \ldots \tau_{N_\tau}\}$ that lie within the range spanning the values of φ_i i.e., (x_1, x_{M^B}). We chose to use equally spaced knots.

Step 5: For each intervention j and function i, we fit a restricted cubic spline regression model to estimate the spline function $\xi_{i,j}(\varphi_i)$, based on the M^B pairs of values $\left\{\overline{\varphi}_{i,m}, \overline{h}_{i,j,m}\right\}$, where $\overline{h}_{i,j,m}$ is the dependent variable, $\overline{\varphi}_{i,m}$ the explanatory variable, and the knots are as chosen in step 4. The estimated spline function $\xi_{i,j}(\varphi_i)$ provides an approximation to $\underset{\varphi^C|\overline{\varphi}_i}{E}\left[h_{i,j}(\varphi^C, \varphi_i)\right]$ that converges to the true expectation as N_θ and M^B increase. Restricted cubic splines can be fitted in many statistical packages; we used the R package

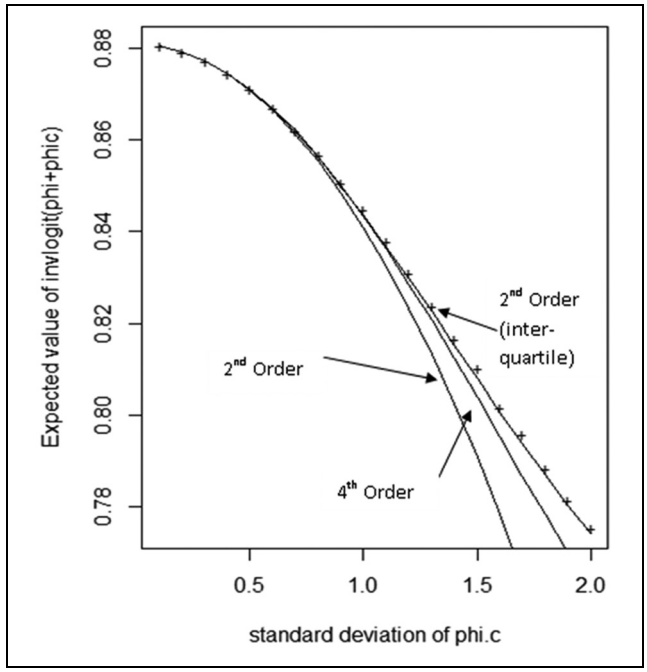

Figure 1 Taylor series approximations for the expectation of the inverse logit of ($\varphi + \varphi_i^C$), where φ is known, φ_i^C is normally distributed, and $E[\varphi + \varphi_i^C] = 2$. Pluses mark true values for the expectation as the standard deviation of φ_i^C varies, and approximations are based around a single expansion around the mean (second- and fourth-order approximation) and averaging over approximations derived at each interquartile mean (second-order approximation).

rms (available at http://biostat.mc.vanderbilt.edu/wiki/Main/Rrms).

EVPPI Calculations for the Fluid Resuscitation Example

We describe below, for the fluid resuscitation example (eq. 1), how EVPPI can be calculated for various different subsets of parameters, representing different potential future studies (Table 3), using the appropriate method described in the Methods section. We assume that we have a set of simulated samples from the joint prior distribution for all of the uncertain parameters.

Example 1: Net Benefit Linear in Nonfocal Parameters

Suppose we wish to evaluate the value of running 2 studies: 1) a 4-arm RCT (such as FEAST) to provide information on relative effects, $\mathrm{d}^{\mathrm{M}}, \mathrm{d}^{\mathrm{S}}$, and 2) a cohort study with 6-mo follow-up to provide information on $\alpha, \mathrm{p}^{\mathrm{B}}, \mathrm{p}^{\mathrm{L}}$. In this case, the focal parameters (i.e., those for which we wish to calculate the EVPPI) are $\varphi = \left\{\mathrm{d}^{\mathrm{M}}, \mathrm{d}^{\mathrm{S}}, \alpha, \mathrm{p}^{\mathrm{B}}, \mathrm{p}^{\mathrm{L}}\right\}$, so that $\varphi^C = \{\mathrm{q}^{\mathrm{M}}, \mathrm{q}^{\mathrm{S}}\}$. By defining

$$f_{1,j}(\varphi) = W\left(\frac{\exp(\alpha)}{(1+\exp(\alpha))} - \frac{\exp(\alpha + d_j^M)}{(1+\exp(\alpha + d_j^M))}\right)$$

$$f_{2,j}(\varphi) = p^L W\left(\frac{\exp(\alpha)}{(1+\exp(\alpha))} - \frac{\exp(\alpha + d_j^M)}{(1+\exp(\alpha + d_j^M))}\right)\left(\frac{p^B \exp(d^S)}{(1+p^B \exp(d^S))}\right)$$

$$f_{3,j}(\varphi) = \left(\frac{\exp(\alpha)}{(1+\exp(\alpha))} - \frac{\exp(\alpha + d_j^M)}{(1+\exp(\alpha + d_j^M))}\right)\left(C^H + p^L C^S \frac{p^B \exp(d^S)}{(1+p^B \exp(d^S))}\right) + C_j^F,$$

the net benefit can be expressed as

$$B_j(\varphi^C, \varphi) = f_{1,j}(\varphi) q^M - f_{2,j}(\varphi) q^S - f_{3,j}(\varphi).$$

This satisfies the conditions that must be met to use method 1 for single-step Monte Carlo estimation of EVPPI. Therefore, we can apply equations 1 and 2 to restate $EVPPI(\varphi)$ in a form requiring only a single simulation step:

$$EVPPI(\varphi) = EVPPI\left(\mathrm{d}^{\mathrm{M}}, \alpha, \mathrm{p}^{\mathrm{B}}, \mathrm{d}^{\mathrm{S}}, \mathrm{p}^{\mathrm{L}}\right) = \underset{\varphi}{E}\left[\max_j \{B_j(\varphi, E[q^M], E[\mathrm{q}^{\mathrm{S}}]) - B_{j*}(\varphi, E[q^M], E[\mathrm{q}^{\mathrm{S}}])\}\right].$$

Method 1 can also be used when $\varphi^C = \{\mathrm{q}^{\mathrm{M}}, \mathrm{p}^{\mathrm{L}}\}$ or when φ^C equals any single element of $\{\mathrm{q}^{\mathrm{M}}, \mathrm{q}^{S}, \mathrm{p}^{\mathrm{L}}\}$.

Example 2: Net Benefit Multilinear in Nonfocal Parameters

Suppose we wish to evaluate the value of running 3 studies: 1) a 4-arm RCT (such as FEAST) to provide information on relative effects, $\mathrm{d}^{\mathrm{M}}, \mathrm{d}^{\mathrm{S}}$, 2) a cohort study with 28-d follow-up (or routine data) to provide information on $\alpha, \mathrm{p}^{\mathrm{B}}$, and 3) a long-term cohort study on patients who survive without NS to provide information on q^M. In this case, $\varphi = \left\{\mathrm{d}^{\mathrm{M}}, \alpha, \mathrm{p}^{\mathrm{B}}, \mathrm{d}^{\mathrm{S}}, q^M\right\}$, so that $\varphi^C = \{p^L, \mathrm{q}^{\mathrm{S}}\}$. Defining

$$f_{1,j}(\varphi) = W\left(\frac{\exp(\alpha)}{(1+\exp(\alpha))} - \frac{\exp(\alpha + d_j^M)}{(1+\exp(\alpha + d_j^M))}\right)\left(\frac{p^B \exp(d^S)}{(1+p^B \exp(d^S))}\right)$$

$$f_{2,j}(\varphi) = C^S\left(\frac{\exp(\alpha)}{(1+\exp(\alpha))} - \frac{\exp(\alpha + d_j^M)}{(1+\exp(\alpha + d_j^M))}\right)\left(\frac{p^B \exp(d^S)}{(1+p^B \exp(d^S))}\right)$$

$$f_{3,j}(\varphi) = \left(\frac{\exp(\alpha)}{(1+\exp(\alpha))} - \frac{\exp(\alpha + d_j^M)}{(1+\exp(\alpha + d_j^M))}\right)\left(Wq^M - C^H\right) - C_j^F$$

net benefit can be expressed as

$$B(j, \varphi^C, \varphi) = -f_{1,j}(\varphi) p^L q^S - f_{2,j}(\varphi) p^L + f_{3,j}(\varphi).$$

This satisfies the conditions that must be met to use method 2 for single-step Monte Carlo estimation of EVPPI. Therefore, we can apply equations 2 and 4 to restate $EVPPI(\varphi)$ in a form requiring only a single simulation step:

$$EVPPI(\varphi) = EVPPI\left(\mathrm{d}^{\mathrm{M}}, \alpha, \mathrm{p}^{\mathrm{B}}, \mathrm{d}^{\mathrm{S}}, \mathrm{q}^{\mathrm{M}}\right) = \underset{\varphi}{E}\left[\max_j \{B_j(\varphi, E[\mathrm{p}^{\mathrm{L}}], E[\mathrm{q}^{\mathrm{S}}]) - B_{j*}(\varphi, E[\mathrm{p}^{\mathrm{L}}], E[\mathrm{q}^{\mathrm{S}}])\}\right].$$

Method 2 can also be used when $\varphi^C = \{q^M, q^S, p^L\}$.

Example 3: Net-Benefit Linear in Functions of the Nonfocal Parameters

Suppose we wish to assess the value of running a cohort study with long-term follow-up, including patients with and without NS, to estimate persistence rates and long-term quality of life.

Here $\varphi = \{q^M, q^S, p^L\}$, so that $\varphi^C = \{d^M, \alpha, p^B, d^S\}$. The net benefit function is no longer multilinear in φ^C, so neither of methods 1 or 2 is valid. However, by defining

$$\beta_{1,j}(\varphi^C) = \frac{\exp(\alpha)}{(1+\exp(\alpha))} - \frac{\exp(\alpha + d_j^M)}{(1+\exp(\alpha + d_j^M))}$$

$$\beta_{2,j}(\varphi^C) = \left(\frac{\exp(\alpha)}{(1+\exp(\alpha))} - \frac{\exp(\alpha + d_j^M)}{(1+\exp(\alpha + d_j^M))}\right) \left(\frac{p^B \exp(d^S)}{(1 + p^B \exp(d^S))}\right)$$

$$f_1(\varphi) = Wq^M - C^H$$

$$f_2(\varphi) = p^L(Wq^S + C^S),$$

the net benefit can be expressed as

$$B_j(\varphi^C, \varphi) = f_1(\varphi)\beta_{1,j}(\varphi^C) - f_2(\varphi)\beta_{2,j}(\varphi^C) - C_j^F.$$

This satisfies the conditions that must be met to use method 3 for single-step Monte Carlo estimation of EVPPI. Therefore, once the unconditional expectations $E[\beta_{1,j}(\varphi^C)]$ and $E[\beta_{2,j}(\varphi^C)]$ are known, we can apply equations 2 and 5 to restate $EVPPI(\varphi)$ in a form requiring only a single simulation step:

$$EVPPI(\varphi) = EVPPI(q^M, q^S, p^L) = \underset{\varphi}{E}\left[\max_j \{B_j(\varphi, E[\beta_{1,j}(\varphi^C)], E[\beta_{2,j}(\varphi^C)]) - B_{j*}(\varphi, E[\beta_{1,j}(\varphi^C)], E[\beta_{2,j}(\varphi^C)])\}\right].$$

Method 3 can also be used when φ is any subset of $\{q^M, q^S, p^L\}$.

Example 4: Net Benefit Includes Joint Nonlinear Functions of Independent Focal and Nonfocal Parameters

Suppose we wish to evaluate the value of running a 4-arm RCT (such as FEAST) to provide information on relative effects, d^M, d^S. Here, $\varphi = \{d^M, d^S\}$, so that $\varphi^C = \{\alpha, p^B, p^L, q^S, q^M\}$. In this case, by defining

$$h_{1,j}(\alpha, \varphi) = \frac{\exp(\alpha + d_j^M)}{(1+\exp(\alpha + d_j^M))}$$

$$h_2(p^B, \varphi) = \left(\frac{\exp(d^S + \ln(p^B))}{(1+\exp(d^S + \ln(p^B)))}\right)$$

$$\beta_1(q^M) = (Wq^M - C^H)$$

$$\beta_2(q^M) = p^L(Wq^S + C^S),$$

the net benefit can be expressed as

$$B_j(\varphi^C, \varphi) = (h_{1,j}(\alpha, \varphi) - h_{1,1}(\alpha, \varphi))(\beta_1(q^M) - h_2(p^B, \varphi)\beta_2(\varphi^C)) - C_j^F.$$

Because α and p^B are independent and d^M and d^S are known, the terms h_1, h_2, β_1, and β_2 are mutually independent. Therefore, we can write

$$EVPPI\left[d^M, d^S\right] = \underset{\varphi}{E}\left[\max_j \{B'_j - B'_{j*}\}\right]$$

where

$$B'_j = B_j(\varphi, \beta_1(E[q^M]), \beta_2(E[p^L], E[q^S]), E[h_{1,,1}(\alpha)], E[h_{1,j}(\varphi, \alpha)], E[h_2(\varphi, p^B)]).$$

$E[h_{1,j}(\varphi, \alpha)]$ (a function of of α and d_j^M) and $E[h_2(\varphi, p^B)]$ (a function of p^B and d^S) can be approximated using Taylor series expansions, as described in the Methods section (method 4). Method 4 can also be used when $\varphi = \{\alpha, p^B\}$.

Example 5: Net Benefit Includes Joint Nonlinear Functions of Correlated Focal and Nonfocal Parameters

Suppose we wish to evaluate the value of collecting routine data on mortality under current practice (i.e., without fluid resuscitation). Here, $\varphi = \{\alpha\}$, so that $\varphi^C = \{p^B, p^L, q^S, q^M, d^M, d^S\}$. In this case, by defining

$$\beta_1(\theta) = \frac{\exp(\alpha)}{(1+\exp(\alpha))}(Wq^M - C^H - p^L\left(\frac{p^B \exp(d^S)}{(1 + p^B \exp(d^S))}\right)(Wq^S + C^S))$$

$$\beta_2(\varphi^C) = p^L\left(\frac{p^B \exp(d^S)}{(1 + p^B \exp(d^S))}\right)(Wq^S + C^S)$$

$$h_{1,j}(\theta) = \left(\frac{\exp(\alpha + d_j^M)}{(1 + \exp(\alpha + d_j^M))}\right)(Wq^M - C^H)$$

$$h_{2,j}(\theta) = \left(\frac{\exp(\alpha + d_j^M)}{(1 + \exp(\alpha + d_j^M))}\right)\left(\frac{p^B \exp(d^S)}{(1 + p^B \exp(d^S))}\right),$$

the net benefit can be expressed as

$$B_j(\varphi^C, \varphi) = \beta_1(\theta) - h_{1,j}(\theta) + h_{2,j}(\theta)\beta_2(\varphi^C) - C_j^F.$$

Because d^M and d^S are correlated, none of the methods previously described can be used to estimate $\underset{\varphi^C}{E}\left[h_{2,j}(\theta)\right]$. However, restricted cubic splines $\xi_j(\alpha)$ can be constructed (method 5) such that

$$\xi_j(\alpha) \simeq \underset{\varphi^C|\alpha}{E}\left[h_{2,j}(\varphi^C, \alpha)\right].$$

Then, EVPPI[$^\alpha$] can be approximated by

$$EVPPI[\alpha] = \underset{\alpha}{E}\left[\max_j\left\{\underset{\varphi^C|\alpha}{E}\left[B_j(\varphi^C, \varphi) - B_{j*}(\varphi^C, \varphi)\right]\right\}\right]$$
$$\simeq \underset{\alpha}{E}\left[\max_j\left\{E\left[\beta_2(\varphi^C)\right](\xi_j(\alpha) - \xi_{j*}(\alpha)) + C_j^F - C_{j*}^F\right\}\right].$$

The method described for this example can also be used when φ is any single element of $\{d^S, d^M, p^B\}$.

RESULTS

Evidence Synthesis and Cost-Effectiveness Analysis

We carried out a Bayesian evidence synthesis, using WinBUGS version 1.4.3,[19] to estimate posterior distributions for the treatment effect parameters d^M and d^S based on the data from the 3 pilot studies (details in the supplementary appendix). A Markov chain Monte Carlo (MCMC) simulation with a burn-in of 50 000 (the number of samples required for convergence of multiple chains) followed by simulation of 5 million parameter value sets (chosen based on the memory limits of the computer) was used to represent joint parameter uncertainty, which forms the prior for the EVPPI analysis.

Table 4 gives posterior parameter means and correlations, based on the pilot studies listed in supplementary appendix A2. The impact of saving a life on the subsequent risk of NS, d^S, is correlated with the effect of each treatment on mortality, d^M. The treatment effects on mortality are also correlated with each other. This is because d^S and d^M are jointly estimated from the pilot studies, where the observed proportion of NS depends on both d^S and d^M (supplementary appendix A2). No other correlations exist between model parameters because they are estimated from independent data sources.

Albumin is the only treatment for which the 95% credible interval for the log-odds ratio relative to no fluids does not include 0 (no effect). Appropriate cost-effectiveness thresholds in an African context are less well established than in regions such as the United Kingdom. We therefore explored a range for the willingness-to-pay threshold (WTP) of 0 to 4000 USD/quality-adjusted life-year (QALY) when calculating net benefit. Based on the pilot data and the synthesis model, the most cost-effective option is albumin if WTP is >$269/QALY and gelofusine otherwise (there are no choices of threshold for which saline or no fluid is the optimal treatment). Around the ''switching'' threshold of $269/QALY, there is considerable uncertainty about the most cost-effective treatment. For example, although gelofusine has the highest expected net benefit at a threshold of $250/QALY, there is only an 18% probability that it is the most cost-effective at that threshold, compared with 54% for albumin, 7% for saline, and 21% for no fluids. Results are more certain at higher thresholds; at $1000/QALY, for example, albumin has the highest net benefit with a 92% probability.

Value of Information

All calculations were performed using the software package R (version 2.15). For the effectiveness parameters d^M and d^S, these samples were generated using MCMC simulation from the evidence synthesis model. For the remaining parameters, samples were generated in R directly. Computations were performed on a PC with an Intel i5 processor and 3 Gb of RAM.

The EVPI is highest around the threshold at which the optimal treatment changes, where the decision is most uncertain and sensitive to additional information (Figure 2). Another influence on the EVPI/WTP relationship is that, at low threshold values, NS costs are high enough for additional information on NS-related parameters (q^S, d^S, p^L, p^B) to potentially change the decision. This effect can be verified by calculating the EVPPI for different model parameters. Table 5 illustrates EVPPI estimates for different subsets of parameters, each calculated in a single simulation step using the methods described above,

Table 4 Results of MCMC Simulation, with Unconditional Means and Correlations between Parameters

Parameter	Description	Posterior Mean and Variance		Correlations								
		Mean	Variance	α	d_2^M	d_3^M	d_4^M	P^B	d^S	p^L	q^M	q^S
α	Baseline mortality (logit scale)	−1.07	0.11	1	0	0	0	0	0	0	0	0
d_2^M	Effect of albumin on mortality	−0.31	0.51	0	1	0.64	0.44	0	0.25	0	0	0
d_3^M	Effect of saline on mortality	−2.34	0.64	0	0.64	1	0.51	0	0.27	0	0	0
d_4^M	Effect of gelofusine on mortality	−0.19	1.81	0	0.44	0.51	1	0	0.27	0	0	0
P^B	Baseline NS	0.10	0.04	0	0	0	0	1	0	0	0	0
d^S	Change to NS risk in ''saved'' patients	1.88	4.52	0	0.25	0.27	0.27	0	1	0	0	0
p^L	Probability that short-term NS proves permanent	0.50	0.08	0	0	0	0	0	0	1	0	0
q^M	QALYs gained by those who survive and are NS-free	19.99	24.97	0	0	0	0	0	0	0	1	0
q^S	QALY loss from NS	5.38	7.94	0	0	0	0	0	0	0	0	1

Note: MCMC = Markov chain Monte Carlo; NS = neurological sequelae; QALY = quality-adjusted life-year.

Table 5 Estimates of EVPPI for a Range of Parameter Subsets, Based on a Willingness-to-Pay Threshold of $250/QALY

Focal Parameter	Method for Single-Step Estimation	EVPPI ($)	Computation Times	
All	NA	561 (EVPI)	Nested 106 × 103 22 s	One Step
dM, α, pB, dS, pL	Method 1	546	26 min	9 s
dM, α, pB, dS, qM	Method 2	415	26 min	10 s
qM	Method 3	73	32 min	21 s
qS	Method 3	0	32 min	25 s
pL	Method 3	239	31 min	20 s
dM, dS	Method 4	342	31 min	1 min 24 s
pB	Method 5	87	31 min	46 s
α	Method 5	0	31 min	44 s
dS	Method 5	243	31 min	46 s
d_2^M(saline)	Method 5	38	32 min	47 s
d_3^M(albumin)	Method 5	$14	32 min	59 s
d_4^M(gelofusine)	Method 5	$24	32 min	46 s

Note: Computation was carried out on a desktop PC with 8 Gb RAM and an Intel i5-2400 processor. EVPI = expected value of perfect information; EVPPI = expected value of perfect parameter information; QALY = quality-adjusted life-year.

assuming a WTP of $250/QALY. Methods that involve replacing nonfocal parameters or functions of them, with their unconditional means in the net benefit function (methods 1–3), can be used to estimate the benefit of knowing individual parameters with certainty in 3 cases: p^L, q^M, and q^S. This benefit is large for the first 2 parameters but 0 for q^S (the QALY loss associated with persistent NS).

The accuracy of EVPPI estimated using methods 1 to 3 depends on the precision of the estimates of the unconditional parameter means, which in turn will depend on the size of the sample generated from the posterior distribution. The point at which sample size is sufficient can be assessed by plotting the relationship between the estimate of EVPPI and the sample size used to derive it. Figure 3 illustrates the number of outer simulations needed for EVPPI estimates to converge in the case where $\varphi = q^M$ and compares nested and single-step simulation. Sample sets of increasing size were used to estimate $EVPPI(\mathrm{q^M})$ using the 1-step approach of method 3 and the 2-step approach with 1000 or 2500 inner simulations for each outer simulation. The (upward) bias resulting from an inadequate number of inner simulations

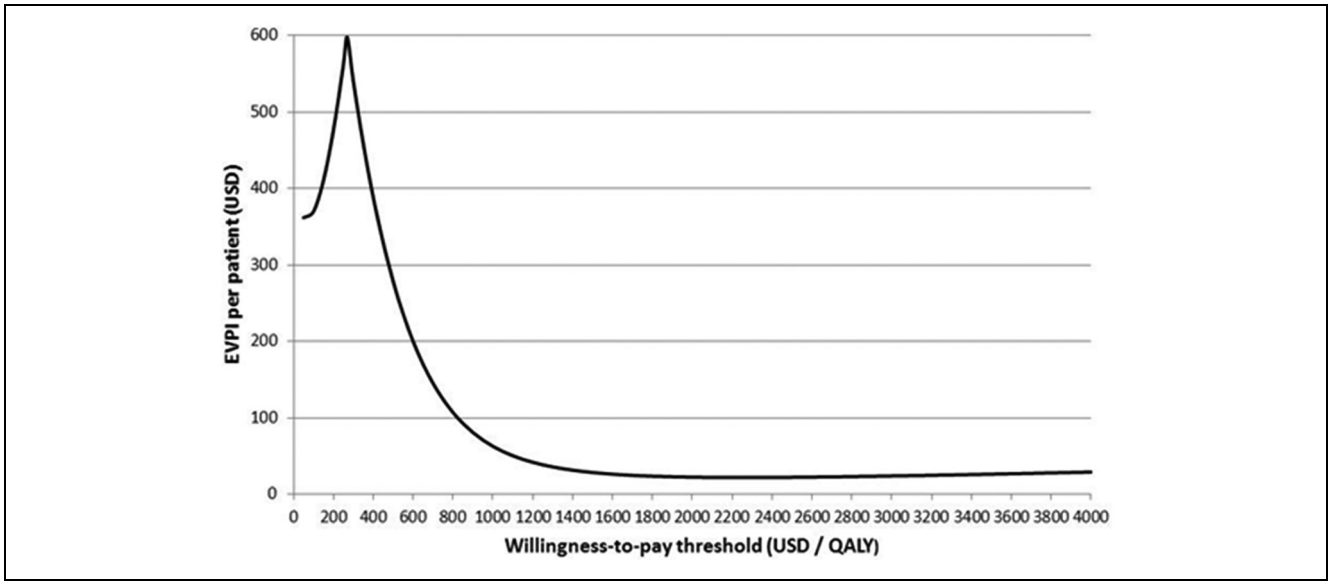

Figure 2 Expected value of perfect information for the fluid resuscitation case study, as a function of the decision maker's willingness-to-pay threshold over the range $50 to $4000 per quality-adjusted life-year.

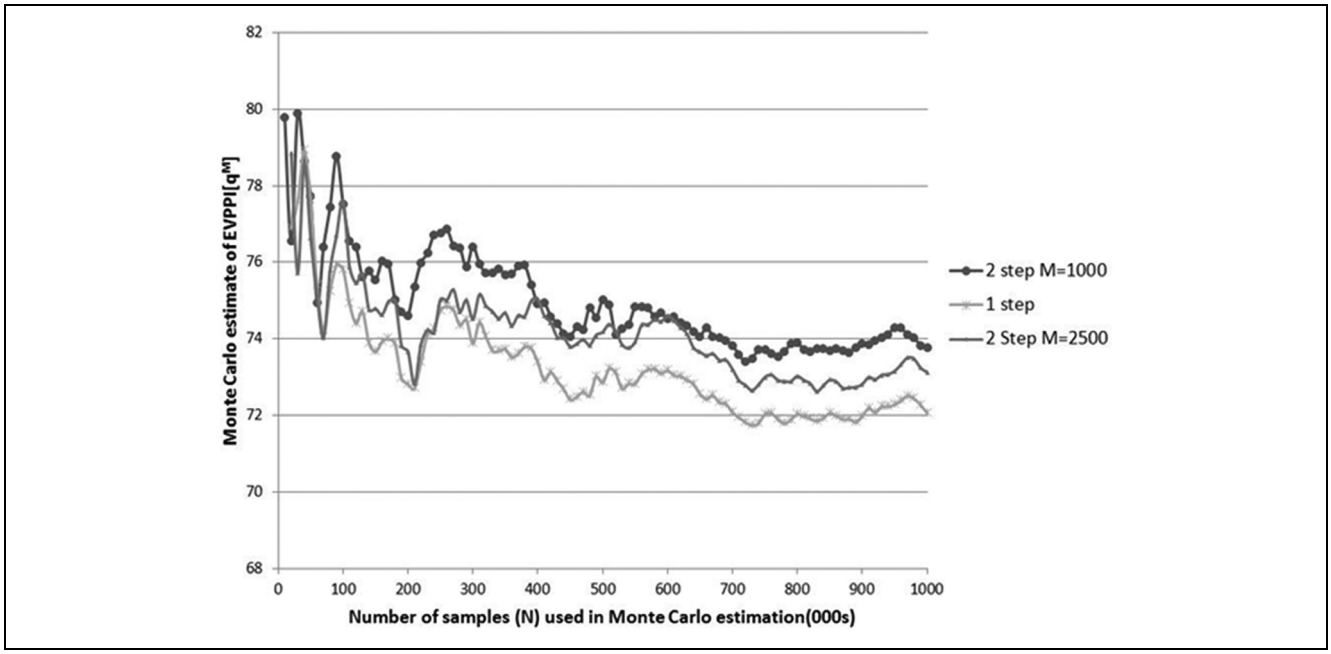

Figure 3 Monte Carlo estimates of the expected value of perfect parameter information of q^M *(quality-adjusted life-years gained through neurological sequelae–free survival) derived using nested versus 1-step Monte Carlo simulation. Estimates were calculated from* N *values sampled from the joint posterior distribution of all parameters. One-step estimation was carried out using method 3. Nested simulation was carried out by subsampling (from the* N *samples)* M *values of the nonfocal parameters* N *times (once for each sampled value of the focal parameter) to estimate conditional expected net benefit.*

can be clearly seen: with 1000 inner samples per outer simulation, the bias is about 3% of the total EVPPI. Nested simulation is also far slower than 1-step simulation; Table 5 gives computation times using each approach, showing the former to be at least 25 times slower when it is based on 1000 simulations

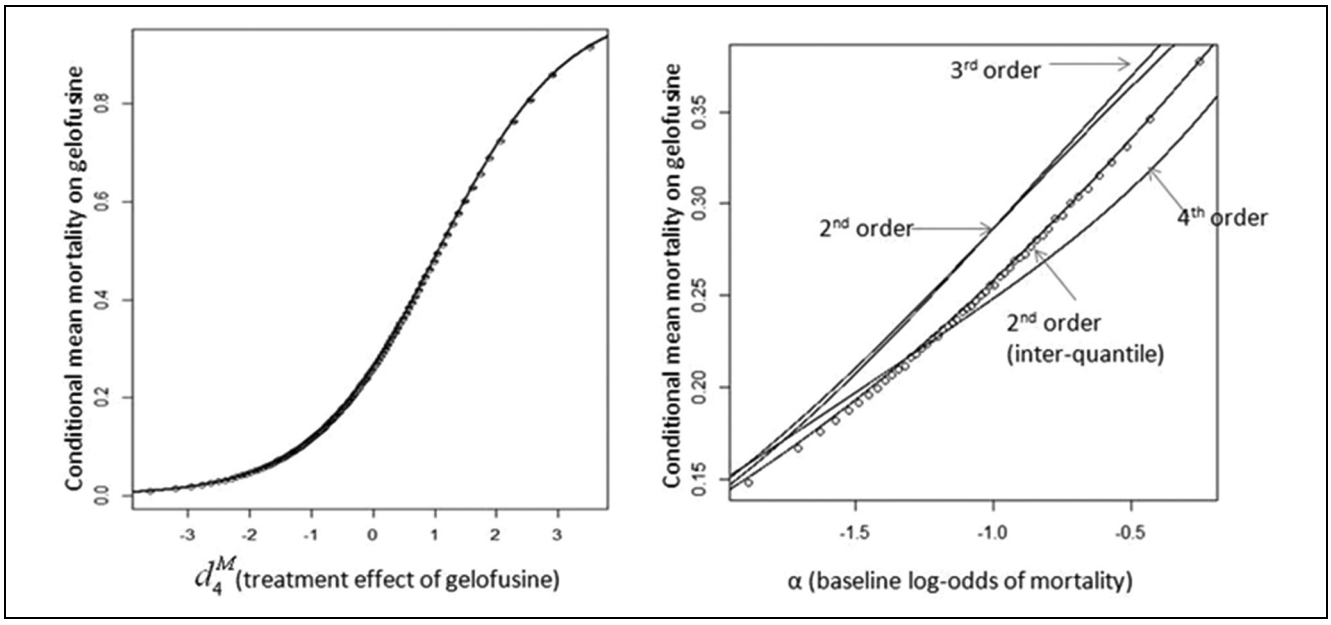

Figure 4 Taylor series approximations to expected mortality on gelofusine conditional on 1) treatment effect of gelofusine (Figure 4.1) and 2) baseline log-odds of mortality (Figure 4.2), for estimation of expected value of perfect parameter information using method 4. Circles represent direct estimates of conditional means, and lines illustrate alternative Taylor series approximations (indistinguishable in the first case).

in the inner loop, and no allowance is made for correlations between parameters (any attempt to incorporate correlations into nested simulation would further increase the speed advantage of the 1-step approach).

Methods 4 and 5 involve constructing functions that give approximate values for conditional means. The accuracy of EVPPI estimates will therefore depend on the goodness of fit of the approximating function, which can be explored graphically. Figure 4 illustrates Taylor series approximations for $E\left[\frac{\exp(\alpha + d_4^M)}{(1 + \exp(\alpha + d_4^M))}\right]$, the expected mortality with gelofusine conditional on perfect information for 1) a known treatment effect d_4^M and 2) a known baseline mortality α. In the first case, the second-order approximation is extremely accurate and cannot be discernibly improved by higher-order terms or the use of interquartile means. In the second case, the second-order approximation is noticeably inaccurate, the fourth-order approximation is still inaccurate but less so, and the approximation based on interquartile means is extremely accurate. However, using the more accurate approximation changes estimated EVPPI by only 1%, from \$345 to \$342 per person (the latter figure, based on the more accurate approximation, is the one quoted in Table 5). Figure 5 illustrates, in the case where $\varphi = d^S$ (example 5), how increasing the number of knots in a spline improves its accuracy, although the difference between 10 and 15 knots is negligible. We found that estimates of EVPPI were extremely stable as the number of knots increased; results in Table 5 were calculated using 10 knots, although results with 6 knots differed by less than 1% in all cases.

DISCUSSION

We have presented a range of methods to avoid the need for nested simulation in EVPPI calculations. The advantages of a 1-step simulation strategy for the computation of EVPPI include unbiased estimation and computational efficiency. It is well known that the inner simulation step can be avoided when the net benefit function is linear or multilinear in the focal parameters (methods 1 and 2).[4,9] We have extended these methods to a broader range of net benefit functions by identifying functions of the focal parameters in which the net benefit function is (multi)linear (method 3), which has been used previously,[17] but the conditions in which the method can be used have not been set out before, and it is not commonly employed. This may be because it

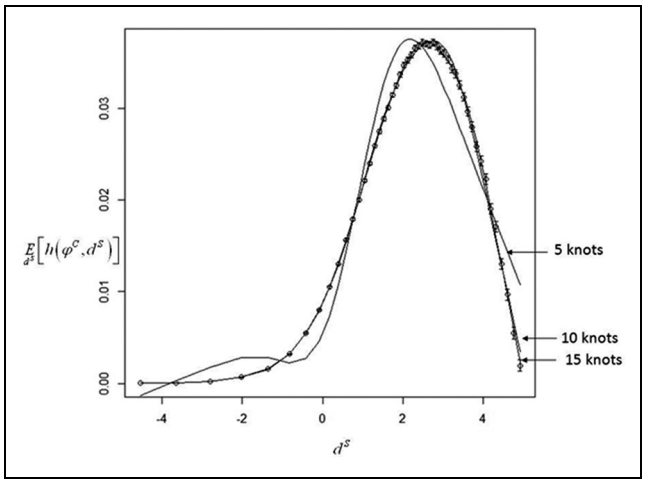

Figure 5 Restricted cubic splines approximating the expectation of $h(\varphi^C, d^S)$ *conditional on* d^S *(impact of successful treatment on risk of neurological sequelae) to allow 1-step estimation of* $EVPPI(d^S)$ *(example 5). Circles represent estimates of the conditional expectation across the plausible range of* d^S*. Error bars represent 95% confidence intervals for each sample mean. The graph shows best-fit splines using 5, 10, and 15 knots.*

requires the net benefit function implied by the economic model to be stated explicitly and to be manipulated algebraically, whereas EVPPI analyses commonly treat the model as a ''black box'' generating net benefit estimates. Ades and others[9] proposed the use of Taylor series approximations to the inverse-logit to approximate conditional expectations. We have set out the situations in which a Taylor series approximation can be used in general (method 4), adapted the approximation for the inverse-logit to improve accuracy, and given the formulae for other commonly used functions. We are unaware of any previous work that has proposed using spline techniques (method 5).

Our use of Taylor series expansions (method 4) and splines (method 5) can be seen as conceptually similar to existing statistical meta-modeling approaches that have been advocated in place of nested simulation, such as Gaussian process emulation.[20] Tappenden and others[21] reviewed potential meta-modeling methods and noted that although they may permit flexible estimation of EVPPI with reasonable accuracy, these methods require substantial specialist expertise. The key difference between meta-modeling and our approximation techniques is that we model components of the net benefit function, rather than the entire economic model. As a result, the methods we present require less technical expertise to implement and review. Furthermore, these methods may be useful for models with a large number of parameters, when Gaussian process emulation can be unfeasible.[7]

The accuracy of approximations to the conditional expectation of the net benefit function should always be checked by comparing approximate with actual values over a plausible range of the focal parameters, as we illustrate in Figure 4. We found that Taylor series expansion around the unconditional mean of the nonfocal parameter could be markedly inaccurate if that parameter had a high variance, even if higher-order terms were included in the expansion. For such situations, we outline an approach averaging over Taylor series expansions around each of the interquartile means, which greatly improves the accuracy of this approach. We found spline regressions were able to provide close approximations to the conditional mean, particularly if enough knots were chosen. The accuracy of spline approximations is largely driven by the accuracy of the bin means to which they are fitted and the coverage of bin means across the plausible range of the focal parameter. In particular, simulated values will be sparse for values of the focal parameter that are less likely. If the optimal treatment changes for these values, it is particularly important to consider the accuracy of the spline approximation of the expectation conditional on these values. The density of conditional mean estimates across the range of the focal parameter can be increased by using more bins, although this will lead to larger sampling errors for each bin mean if the total number of samples is kept constant. The number of samples N_θ used for spline estimation should be large enough to support enough bins across the plausible range of the focal parameter while ensuring sufficiently low standard errors for each bin mean.

A particular challenge for EVPPI calculations occurs when correlations exist between components of φ^C and components of φ.[5] In this situation, values at each inner simulation need to be sampled from the conditional distribution of φ^C; otherwise, EVPPI estimates will be downwardly biased. One approach would be to carry out a separate MCMC simulation for each realization from the outer simulation. This will give samples of φ^C from the correct distribution but at considerable computational expense. A more efficient approach may be to assume a parametric joint distribution for the correlated parameters, from which a conditional sampling distribution can be derived at each inner simulation. However, this may introduce an additional source of bias. One-step estimation of EVPPI using spline approximations allows us to side step this problem.

The purpose of VoI is to guide decisions on research prioritization, so that it is necessarily conducted in situations in which current information is inadequate and assumptions may be proven quite wrong by later research. The goal is to reflect as fairly as possible current information, however sparse that might be. The credibility of VoI estimates therefore depends on the credibility of the cost-effectiveness model on which they are based and how accurately prior distributions reflect prestudy parameter uncertainty, which requires guidance from clinicians and decision makers regarding assumptions made in the model and plausible alternatives. Decision makers might require the model to be extended to allow for different severities of NS or the health utility loss due to short-term NS to be included. We have assumed fixed treatment effects across studies and used pilot study data to estimate them. These data were the only evidence available to us and were insufficient to estimate less restrictive models; thus, it is possible that we have underestimated uncertainty in treatment effect estimates. Furthermore, use of pilot study data may introduce bias, as their design and conduct may not be as rigorous as a full RCT. Very little information was available on the long-term costs and health consequences of NS, and the values chosen for this analysis were therefore not evidence based but chosen to ensure nonzero EVPPI estimates for these parameters. When existing evidence is sparse and/or of poor quality, as with long-term NS in our case study, it is important to be able to somehow characterize the extent of the uncertainty, as VOI will be highly sensitive to this. One possibility is to incorporate formal expert elicitation of plausible parameter values, together with opinion on the possible degree of bias in the existing evidence base.[22] Finally, we would advocate the use of sensitivity analyses to explore the sensitivity of VoI results to modeling assumptions,[17] as would be conducted as standard practice for cost-effectiveness analysis.[23]

The focus of this article is to illustrate methods for 1-step estimation of EVPPI. Because of the limitations set out above, we do not intend to draw conclusions about research priorities in the area of fluid resuscitation. Nevertheless, the case study illustrates how EVPPI calculations can be informative when considering the relative merits of efficacy trials compared with longer-term epidemiological studies or when selecting outcome measures and treatments to include in a proposed trial. For example, there was considerable debate following the pilot studies around the inclusion of albumin in the FEAST trial, driven by perceptions that this treatment was too costly for use in Africa. Our EVPPI analysis suggests that there was considerable value to including this arm, although the caveats mentioned above apply. The results from the FEAST trial have recently been published,[16] with robust but unexpected findings (harm was demonstrated in the fluid resuscitation arms).

A key limitation of our approach is that the methods rely on being able to explicitly state the net benefit function implied by the model. We would argue that there are clear benefits in explicitly setting out the net benefit function, in terms of transparency, understanding, and communication of assumptions, as well as the reduced computational burden for VoI calculations, and would always recommend doing this unless the decision model is unavoidably too complex to allow it. We have presented the use of univariate splines in our case study, which can be used only if there are no nonlinear functions of multiple correlated focal parameters in the net benefit function. When this is not the case (for example, if we were to consider 2- or 3-arm trials in the fluid replacement example), then the spline approach (method 5) could be implemented through the use of multivariate adaptive regression splines.[7]

In this article, we have focused on EVPPI; however, the methods presented here can also be used in the calculation of the expected value of sample information (EVSI). EVSI is more computationally intensive than EVPPI because it requires an extra layer of simulation required to simulate new data from a given new study design. Furthermore, if we wish to optimize over potential new study design factors, then EVSI will need to be evaluated repeatedly, which will substantially increase the computational burden. Ades and others[9] set out algorithms for the estimation of EVSI in a number of situations, all of which require nested Monte Carlo simulation unless the net benefit function satisfies the conditions required for use of methods 1 and 2. The methods presented here extend the circumstances in which we can avoid the inner simulation step for EVSI, hence improving computing times and thereby making the use of EVSI feasible in a wider range of situations.

REFERENCES

1. Schlaifer R. Probability and Statistics for Business Decisions: An Introduction to Managerial Economics under Uncertainty. New York: McGraw-Hill; 1959.

2. Raïffa H, Schlaifer R. Applied Statistical Decision Theory. New York: Wiley; 2000.

3. Gelman A, Carlin JB, Stern HS, Rubin DB. Bayesian Data Analysis. 2nd ed. New York: Taylor & Francis; 2003.

4. Felli JC, Hazen GB. Sensitivity analysis and the expected value of perfect information. Med Decis Making. 1998;18(1):95–109.

5. Thompson KM, Evans JS. The value of improved national exposure information for perchloroethylene (Perc): a case study for dry cleaners. Risk Anal. 1997;17(2):253–71.

6. Sutton AJ, Welton NJ, Cooper NJ, Abrams KR, Ades AE. Evidence Synthesis for Decision Making in Healthcare. New York: Wiley; 2012.

7. Brennan A, Kharroubi S, O'Hagan A, Chilcott J. Calculating partial expected value of perfect information via Monte Carlo sampling algorithms. Med Decis Making. 2007;27(4):448–70.

8. Oakley JE, Brennan A, Tappenden P, Chilcott J. Simulation sample sizes for Monte Carlo partial EVPI calculations. J Health Econ. 2010;29(3):468–77.

9. Ades AE, Lu G, Claxton K. Expected value of sample information calculations in medical decision modeling. Med Decis Making. 2004;24(2):207–27.

10. Dondorp AM, Fanello CI, Hendriksen IC, et al. Artesunate versus quinine in the treatment of severe falciparum malaria in African children (AQUAMAT): an open-label, randomised trial. Lancet. 2010;376(9753):1647–57.

11. Brewster DR, Kwiatkowski D, White NJ. Neurological sequelae of cerebral malaria in children. Lancet. 1990;336(8722):1039–43.

12. Mung'Ala-Odera V, Snow RW, Newton CR. The burden of the neurocognitive impairment associated with *Plasmodium falciparum* malaria in Sub-Saharan Africa. Am J Trop Med Hygiene. 2004;71(2 suppl):64–70.

13. Maitland K, Pamba A, Newton CRJC, Levin M. Response to volume resuscitation in children with severe malaria. Pediatr Crit Care Med. 2003;4(4):426–31.

14. Maitland K, Pamba A, English M, et al. Pre-transfusion management of children with severe malarial anaemia: a randomised controlled trial of intravascular volume expansion. Br J Haematol. 2005;128(3):393–400.

15. Akech S, Gwer S, Richard I, et al. Volume expansion with albumin compared to gelofusine in children with severe malaria: results of a controlled trial. Plos Clin Trials. 2006;1(5):e21.

16. Maitland K, Kiguli S, Opoka RO, et al. Mortality after fluid bolus in African children with severe infection. N Engl J Med. 2011;364(26):2483–95.

17. Welton NJ, Ades AE, Caldwell DM, Peters TJ. Research prioritization based on expected value of partial perfect information: a case-study on interventions to increase uptake of breast cancer screening. J R Stat Soc A Stat. 2008;171:807–34.

18. Harrell FE. Regression Modelling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis. New York: Springer; 2010.

19. Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS: a Bayesian modelling framework: concepts, structure, and extensibility. Stat Comput. 2000;10(4):325–37.

20. Stevenson MD, Oakley J, Chilcott JB. Gaussian process modeling in conjunction with individual patient simulation modeling: a case study describing the calculation of cost-effectiveness ratios for the treatment of established osteoporosis. Med Decis Making. 2004;24(1):89–100.

21. Tappenden P, Chilcott JB, Eggington S, Oakley J, McCabe C. Methods for expected value of information analysis in complex health economic models: developments on the health economics of interferon-beta and glatiramer acetate for multiple sclerosis. Health Technol Asses. 2004;8(27):1–78.

22. Turner RM, Spiegelhalter DJ, Smith GCS, Thompson SG. Bias modelling in evidence synthesis. J R Stat Soc Ser A Stat Soc. 2009; 172(1):21–47.

23. Caro JJ, Briggs AH, Siebert U, Kuntz KM. Modeling good research practices: overview. A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force–1. Med Decis Making. 2012;32(5):667–77.