Value of Information Analysis in Models to inform Health Policy

Christopher H. Jackson, Gianluca Baio, Anna Heath, Mark Strong, Nicky J. Welton, Edward C. F. Wilson

xxxxxx 0000. 00:1–26

Keywords

decision theory, evidence synthesis, health economics, Bayesian, sensitivity analysis, design

Abstract

Value of Information (VoI) is a decision-theoretic approach to estimating the expected benefits from collecting further information of different kinds, in scientific problems based on combining one or more sources of data. VoI methods can assess the sensitivity of models to different sources of uncertainty, and help to set priorities for further data collection. They have been widely applied in healthcare policy making, but the ideas are general to a range of evidence synthesis and decision problems. This paper will give a broad overview of VoI methods, explaining the principles behind them, the range of problems that can be tackled with them, and how they can be implemented, and discuss the ongoing challenges in the area.

Contents

1. Introduction

Many scientific investigations involve making decisions or inferences using a model that combines more than one piece of information. Typically the conclusions are subject to uncertainty. It is then natural to ask which information is the weakest. Specifically we want to know which contributes most to uncertainty about our conclusion (noting that not all weak information might affect the conclusion). These kinds of questions are addressed by methods for *sensitivity analysis*. However, suppose that the question were expressed in a slightly different way, to ask *"on which part of the model would further information be most valuable?"*. Then we might ask what value there would be in collecting a particular number of observations from a study of a particular design. These questions are directly answered by *Value of Information* (VoI) methods, a set of tools that encompass sensitivity analysis, research prioritisation and design of experiments. This article gives a broad overview of the principles of VoI methods and the range of applications they have been used in, and discusses the ongoing challenges.

Value of Information analysis is based on Bayesian decision theory, where a model with parameters $\boldsymbol{\theta}$ is used as the basis of some action d. Uncertainty about $\boldsymbol{\theta}$ is quantified by a (joint) probability distribution, we quantify the losses $L(d, \boldsymbol{\theta})$ associated with different actions, and choose the action that minimises the expected loss. In some situations we might express the problem in terms of *maximising* the expected *benefit*, but this is mathematically equivalent. This framework can describe both "decision" problems where we choose from a finite set of actions, and statistical inference problems such as parameter estimation.

A key concept is the expected value of perfect information (EVPI) (Raiffa & Schlaifer 1961), which can be framed as the expected value of eliminating all uncertainty about $\boldsymbol{\theta}$, or alternatively, as the expected loss due to the uncertainty. A second key concept is the expected value of sample information (EVSI), which quantifies the value of undertaking a specific data collection exercise in order to reduce uncertainty. Bayesian methods for

experimental design (Lindley 1956) can be framed as computations of the EVSI in which the decision problem is estimation of some unknown quantity (Raiffa & Schlaifer 1961).

The VoI approach has been investigated in depth as a way of prioritising further research in the context of health policy-making (Felli & Hazen 1998, Willan & Pinto 2005, Claxton & Sculpher 2006, Welton et al. 2008, Wilson 2015). This is because in many health policy-making settings, decisions on treatment funding are informed by synthesis of evidence of effects and costs, under a model where parameter uncertainty is quantified probabilistically (see, e.g. Claxton et al. 2001, Welton et al. 2020). An international network has recently been founded to promote and develop VoI methods in this area: `http://convoi-group.org`. VoI methods have also been applied in other areas of decision-making, e.g. engineering (Davis et al. 1972, Eidsvik et al. 2015) and management (Repo 1989).

Health economic models are usually designed to inform a decision between a finite number of policies. However, the ideas of VoI are applicable to other kinds of modelling questions. For example, *mathematical* or *computer* models are typically designed to estimate one or more (unobservable) quantities defined as a function of uncertain inputs, without necessarily having an associated decision. While sensitivity analysis in these models is a much-studied field (Saltelli et al. 2000), the connection with research prioritisation and study design is less well-known (Borgonovo et al. 2016). Jackson et al. (2019) demonstrated the use of VoI in an even broader class of models known as Bayesian *multiparameter evidence synthesis*. In these models, which have been used in epidemiology and public health (Welton et al. 2012, De Angelis et al. 2014), evidence sources are related stochastically and deterministically through a directed acyclic graph, and the purpose of modelling is to estimate one or more quantities of interest.

In this article, the VoI measures are defined in Section 2 in their broadest generality. Section 3 explains how these can be defined and applied in decision problems, with illustrations from health economic modelling, and Section 4 explains how VoI principles can be used for sensitivity analysis and research design in estimation problems, with examples from models to inform public health. Section 5 reviews computation of VoI measures, a major challenge of this field, detailing several recently-developed methods that were motivated by the demands of healthcare decision models. A final discussion summarises some general challenges.

Value of Information methods: overview

- Value of Information (VoI) methods are a Bayesian decision-theoretic approach to sensitivity analysis and research prioritisation.
- VoI measures describe expected loss reductions (or benefit gains) from collecting further information.
- They can be used in models designed either for decision making or estimation. Uncertainty in model parameters must be described by probability distributions.
- For decision making, the model output must describe the losses or benefits associated with different decisions, given different parameter values.
- VoI ideas can also be used in models for estimating a quantity of interest, where gains from information are described by reductions in the variance of the estimate.
- The expected value of perfect information (EVPI) describes the value of eliminating all uncertainty.
- The expected value of partial perfect information (EVPPI) describes the value of

eliminating all uncertainty about a particular parameter or group of parameters, thus measures sensitivity of the decision or estimate to the parameter(s).
- The expected value of sample information (EVSI) describes the benefits from a study of a particular design and sample size.
- The expected net benefit of sampling (ENBS) is the EVSI of a particular study minus its cost, a measure of the expected return on investment from that study.

2. Formal definitions

To use Value of Information methods, it is required to take a Bayesian decision-theoretic perspective. We have a model with parameters $\boldsymbol{\theta}$, where uncertainty about $\boldsymbol{\theta}$ can be described by a joint probability distribution $p(\boldsymbol{\theta})$. This distribution describes beliefs under *current information*, which could be viewed as the posterior distribution after all currently-available data have been observed, or as the prior before any planned future data are collected. Most generally, the purpose of the model is to inform a decision or action d from a space of possible decisions $\mathcal{D}$, to minimise an expected loss $E_{\boldsymbol{\theta}}(L(d, \boldsymbol{\theta}))$ with the expectation taken with respect to the distribution of $\boldsymbol{\theta}$.

In following sections, we explain that through specific definitions of the decision space and the loss function, this perspective can deal with situations where the purpose of the model is either:

- to make a decision between finite set of actions (Section 3), or
- to estimate one or more quantities of interest (Section 4)

In either case, Value of Information methods describe the expected benefits of further information in terms of *reductions in expected loss*. Note that in the estimation framework, the action is framed as minimising a loss, whereas in the finite-decision framework we talk about maximising a benefit, but these perspectives are equivalent if the benefit is defined as minus the loss.

The key VoI measures are defined as follows, in each instance, as the difference between the loss $E_{\boldsymbol{\theta}}(L(d^*, \boldsymbol{\theta}))$ of the decision under *current* information, and the expected loss of the decision given *better* information.

1. The *expected value of perfect information* (EVPI) is the expected loss of the optimal decision $d^* = \arg\min_d E_{\boldsymbol{\theta}}\{L(d, \theta)\}$ made under current information, minus the expected loss for the decision $d^*_{\boldsymbol{\theta}} = \arg\min_d L(d, \boldsymbol{\theta})$ we would make if we knew the true $\boldsymbol{\theta}$ (Raiffa & Schlaifer 1961).

$$EVPI = E_{\boldsymbol{\theta}}\{L(d^*, \boldsymbol{\theta})\} - E_{\boldsymbol{\theta}}\{L(d^*_{\boldsymbol{\theta}}, \boldsymbol{\theta})\} \quad 1.$$

   This describes the overall value of the *decision uncertainty*, and gives an upper bound on the expected gains from any new information (Parmigiani & Inoue 2009).
   Another interpretation is as the expectation of the "opportunity loss" $L(d^*, \boldsymbol{\theta}) - L(d^*_{\boldsymbol{\theta}}, \boldsymbol{\theta})$ that is, what we expect to lose due to imperfect information.
2. The *expected value of partial perfect information* (EVPPI) for a particular parameter $\boldsymbol{\phi}$ is the expected reduction in loss if $\boldsymbol{\phi}$ were to be learnt precisely. Since this precise

value is not yet known, an expectation must be taken over all possible values.

$$EVPPI_{\boldsymbol{\phi}} = E_{\boldsymbol{\theta}}\{L(d^*, \boldsymbol{\theta})\} - E_{\boldsymbol{\phi}}[E_{\boldsymbol{\theta}|\boldsymbol{\phi}}\{L(d^*_{\boldsymbol{\phi}}, \boldsymbol{\theta})\}] \quad 2.$$

where $d^*_{\boldsymbol{\phi}} = \arg\min_d E_{\boldsymbol{\theta}|\phi}\{L(d, \theta)\}$ is the optimal decision we would make if $\boldsymbol{\phi}$ were known. This is an upper bound on the potential value of data $\mathbf{y}$ which inform only $\boldsymbol{\phi}$ (i.e that are conditionally independent of $\boldsymbol{\theta}$ given $\boldsymbol{\phi}$).

3. The *expected value of sample information* $EVSI_{\mathbf{y}}$ is the reduction in loss we would expect from collecting an additional dataset $\mathbf{y}$ of a specific design and making a decision $d^*_{\mathbf{y}} = \arg\min_d E_{\boldsymbol{\theta}|\mathbf{y}}\{L(d, \theta)\}$:

$$EVSI_{\mathbf{y}} = E_{\boldsymbol{\theta}}\{L(d^*, \boldsymbol{\theta})\} - E_{\mathbf{y}}\left[E_{\boldsymbol{\theta}|\mathbf{y}}\{L(d^*_{\mathbf{y}}, \boldsymbol{\theta})\}\right] \quad 3.$$

The inner expectation is now with respect to the (as yet unknown) posterior distribution of $\boldsymbol{\theta}|\mathbf{y}$, after learning $\mathbf{y}$.

4. The *expected net benefit of sampling* for an dataset $\mathbf{y}$ is $ENBS_{\mathbf{y}} = EVSI_{\mathbf{y}} - C_{\mathbf{y}}$, where $C_{\mathbf{y}}$ measures the costs of obtaining the data, measured on the same scale as the loss or VoI. Typically both the cost and the EVSI would increase with the sample size, and we might choose the optimal sample size (or other aspects of the design) to maximise $ENBS_{\mathbf{y}}$.

The following two sections will give examples that contrast the complementary goals of VoI analysis: to compare the sensitivity of a decision or estimate to different sources of uncertainty, and to guide and design further data collection to reduce uncertainty.

3. Health economic and related decision problems

This section describes the most common application of Value of Information methods, to *health economic evaluations* based on decision models. The principles of VoI described in this section are conceptually the same for any situation where a model is used to represent a decision between alternative policies, and parameter uncertainties are described with probability distributions. A brief overview of health economic modelling is given here — for a fuller review from a statistical perspective see, e.g. Baio (2018), Welton et al. (2020).

In a typical health economic decision model, a policy-maker compares D alternative health policies, and the optimal policy is the one that maximises the *expected net benefit* $E_{\boldsymbol{\theta}}(NB_d(\boldsymbol{\theta}))$. For example, with $D = 2$, a novel treatment $d = 1$ might be compared to current best practice or "standard of care" $d = 0$. The net benefit combines health outcomes, costs and the decision-maker's relative valuation of these two quantities. A typical definition of this is a linear function, $NB_d(\boldsymbol{\theta}) = B_d(\boldsymbol{\theta})\lambda - C_d(\boldsymbol{\theta})$ (net monetary benefit) where $B_d(\boldsymbol{\theta})$ is the expectation (over individuals) of a health outcome (commonly *quality-adjusted life years* (QALY), a measure combining overall survival and level of health) for people treated under policy d, $C_d(\boldsymbol{\theta})$ is the corresponding expected cost, and λ is the decision-maker's *willingness to pay* for one unit of health benefit. The functions $B_d(\boldsymbol{\theta}), C_d(\boldsymbol{\theta})$ can be constructed in many ways (Brennan et al. 2006). A simple decision tree model is illustrated in Baio (2018), and another common approach uses a Markov or similar *state-transition* model, where states of health or clinical events are associated with distinct health values and costs, and expected health or cost outcomes are calculated with respect to probabilities of following particular pathways through states. The model parameters $\boldsymbol{\theta}$ usually include, for example,

the effectiveness of treatments, and the expected costs and probabilities of experiencing state transitions or clinical events, and their expected values and uncertainty distributions are informed from all relevant evidence.

3.1. Use of VoI in health economic modelling

As the parameters are uncertain, there is a chance that the decision may not be the best one. Nevertheless, a risk-neutral decision maker seeking to maximise overall population benefits (Arrow & Lind 1970) should make the decision that maximises the expected net benefit under current information, whatever the level of uncertainty (Claxton 1999). However, if further information is collected, this may lead to a different optimal decision, hence improvements in net benefit. Value of Information analysis can quantify these expected improvements from further information. This allows the decision maker to supplement their decision (under current information) with a recommendation on whether further data collection would be worthwhile. The decision maker might also approve an intervention for use only in the context of research (Walker et al. 2012).

Defining the loss as $L(d, \boldsymbol{\theta}) = -NB_d(\boldsymbol{\theta})$ we can express the decision problem as maximising the expected net benefit, rather than minimising the expected loss. We then rewrite expressions 1–3, to explicitly define what each VoI measure looks like for decision problems such as these which are defined as a choice between a finite number of actions.

$$EVPI = \max_d\{E_{\boldsymbol{\theta}}(NB_d(\boldsymbol{\theta}))\} - E_{\boldsymbol{\theta}} \max_d\{NB_d(\boldsymbol{\theta})\}$$

This can be converted to a "population-level" VoI by multiplying by the number of people would would benefit from the decision, now and in the future, over the time horizon that the decision would remain relevant for, discounted to present value (Philips et al. 2008). If the population-level expected value of perfect information (EVPI) (using net monetary benefit) is less than the cost of further research, then further research of any kind is judged to be not worthwhile. Otherwise, we can examine the expected value of partial perfect information (EVPPI) for each parameter $\boldsymbol{\phi}$, or groups of parameters jointly. Again research is judged to be not worthwhile for parameters whose population EVPPI is less than the cost of research.

$$EVPPI(\boldsymbol{\phi}) = \max_d\{E_{\boldsymbol{\theta}}(NB_d(\boldsymbol{\theta}))\} - E_{\boldsymbol{\phi}} \max_d\{E_{\boldsymbol{\theta}|\boldsymbol{\phi}}(NB_d(\boldsymbol{\theta}))\} \quad 4.$$

Then we might consider proposing a specific study to obtain data that would inform the quantities $\boldsymbol{\phi}$ where research might be valuable. The EVSI measures the expected health benefits from the information $\mathbf{y}$ from this study.

$$EVSI(\mathbf{y}) = \max_d\{E_{\boldsymbol{\theta}}(NB_d(\boldsymbol{\theta}))\} - E_{\mathbf{y}} \max_d\{E_{\boldsymbol{\theta}|\mathbf{y}}(NB_d(\boldsymbol{\theta}))\} \quad 5.$$

Finally we could compare the costs and benefits of different study designs, and recommend a design or sample size that maximises the expected net benefit of sampling.

Note that the EVPI exceeding the cost of research is necessary but not sufficient for further research to be worthwhile — a *sufficient* condition is given by the EVSI exceeding the cost of research. Note also that the use of VoI assumes that the space of possible decisions, and how they are are valued, will be the same after new information emerges. In practice, for example, new treatments may become available in the future, though this is unknowable in advance.

3.2. VoI used in practice: a case study and example workflow

The following box describes a real health economic evaluation where a VoI analysis was conducted to inform research prioritisation. Reeves et al. (2019) compared three strategies for dressing wounds after surgery to reduce surgical site infection (SSI). An EVPPI and EVSI analysis provided evidence that a large trial to provide better information about dressing effectiveness would lead to substantially better-informed policy on use of health service resources. Specifically, the trial would provide population health benefits that would be at least worth the cost of the trial. The box illustrates a typical "workflow" that is followed, showing the steps of problem definition, evidence gathering, model building and analysis that are involved in this kind of application of VoI in decision modelling.

Example workflow for a VoI calculation in a health economic model, from Reeves et al. (2019)

1. **Define the decision problem**: compare cost-effectiveness of three strategies for surgical wound dressings to reduce surgical site infection (SSI), from the perspective of the health service.
2. **Identify relevant evidence** from systematic review, in conjunction with **developing a model** to estimate outcomes of interest to the decision-maker.
   Evidence on the intervention effectiveness is usually the most important. Here, summary estimates from network meta-analysis indicated an odds ratio of infection of 1.05 (95% credible interval 0.37 to 2.41) for glue compared to simple dressings, and 0.98 (0.56 to 1.55) for no dressing compared to simple, showing substantial uncertainty.
   Here the model estimates net monetary benefit of the interventions by combining the effectiveness data with data on the baseline risks of SSI, the costs of dressing, SSI management and other healthcare, and the quality-of-life impact of SSI.
3. **Estimate EVPI and EVPPI** per patient, and convert to a population level using estimates of the number of people who would benefit from the policy over the "lifetime" that it would be used for. In this example, this was the population with surgery wounds, and a "lifetime" of five years. The parameters with the greatest EVPPI were those describing the relative effectiveness of the dressing strategies, with EVPPI of around £2000 million, much greater than the cost of a trial.
4. **Estimate EVSI** to inform the design of a study to give evidence about any parameters with substantial EVPPI. In this example, various designs were compared for a randomised controlled trial of the three strategies, with different sample sizes and randomisation ratios.
5. **Gather further evidence if possible**, informed by the EVSI calculation. In this example, a "pilot" study of 394 patients was also conducted, both to obtain a small amount of further evidence, and to inform the feasibility of a larger study. The EVSI is calculated with and without the data observed from this study, which when added to the meta-analysis, gave slightly more precise estimates of the odds ratios of SSI: 0.98 (0.48 to 1.75) for glue compared to simple dressings and 1.02 (0.64 to 1.49) for exposed compared to simple dressings.
6. **Make further research recommendation.** Before the pilot data, the population EVSIs for larger studies of 3000 patients were around £2000 million, and afterwards

around £1500 million, indicating substantial value in a large study that would vastly exceed the cost of conducting it. The optimal design might also be chosen to minimise the ENBS, though this was not considered in the published example.

3.3. Open challenges for VoI in health economic modelling

Here we give a brief overview of practical challenges and current issues relating to VoI analysis in health economic modelling. Challenges with *computation* of VoI quantities are discussed later, in Section 5.

Firstly, a VoI analysis requires specifying the decision problem and outcomes of interest, constructing a model to represent the problem, choosing relevant sources of evidence, and quantifying parameters in the model. There are guidelines for good practice in general (Caro et al. 2012), and for specific settings (e.g. in the U.K., National Institute for Health and Care Excellence (2013)). The guidelines highlight challenges in specific parts of this process, and many of these will affect estimates of VoI (Koffijberg et al. 2018), though these are challenges of health economic evaluation in general. As VoI is explicitly about quantifying decision uncertainty, then one particular issue pertinent to VoI is how uncertainty about the knowledge relevant to the decision should be quantified. Many kinds of uncertainty are difficult to represent as parameters in models, an issue sometimes described as "structural" uncertainty. This can include not only statistical model uncertainty (about how to describe a specific dataset) but also uncertainties about how to represent the clinical and economic process, or what data are relevant. For such uncertainties, it is typical to present results under different assumptions or scenarios (Briggs et al. 2012). This can aid decision making, together with qualitative judgments about which assumptions are more plausible. However, presented in this way, it is hard to determine the value of further information intended to resolve uncertainty about which scenario should be preferred. Methods for quantifying "structural" uncertainies have generally involved parameterising these uncertainties within a single model, and using evidence elicited from experts or information about statistical goodness-of-fit to obtain distributions for those parameters (Jackson et al. 2011, Strong & Oakley 2014, Thom et al. 2017). This enables estimation of the value of partial perfect information about the structural uncertainty. However, such evidence may be weak, and it may be difficult to construct a single encompassing model — the range of "structural" uncertainties that can be conceived is as wide as the range of models that can be conceived. It may also be difficult to obtain data to reduce the uncertainty. More broadly, this is an example of a common dilemma in statistical sciences — given weak evidence, is it better to quantify judgements, or just to present different evidence qualitatively? Practical guidance for particular situations in health economic modelling would help to make VoI methods more useful in practice.

4. Value of Information in estimation problems

Mathematical models representing complex real-world processes, based on evidence from different sources, are often used to estimate some quantity or quantities that are not directly observable. While the estimates might be of interest to policy-makers, the model might not describe a specific policy decision. Value of Information ideas can still be used to determine

which of the uncertain model inputs contribute most to uncertainty about the model output (*sensitivity analysis*). Then, determining what specific further data would be most valuable is the domain of *experimental design*, which, when done from a Bayesian perspective, is a form of Value of Information analysis.

The methods described here can be used in two important classes of models, though the application of VoI is similar in both:

1. *computer models* (or "mathematical" models) where the output is a known deterministic function of the inputs (Oakley & O'Hagan 2004), $\alpha = g(\boldsymbol{\theta})$. This can include stochastic models, such as "microsimulation" models that simulate random outcomes for individuals, if the "output" is defined as an average over a sufficiently large number of random simulations.
2. *Bayesian multiparameter evidence synthesis* (Welton et al. 2012), where unknown parameters $\boldsymbol{\theta}$ and observed data are related through a directed acyclic graph. The model is designed to estimate a subset or function of $\boldsymbol{\theta}$, denoted α.

Sections 4.4.1 and 4.4.2 describe examples of how these kinds of models have been used in health policy.

Methods for sensitivity analysis in computer models have been widely investigated, since the relationship of the input to the output is typically complex. See, for example, Saltelli et al. (2000), or Borgonovo & Plischke (2016) and Razavi et al. (2021) for more recent reviews. A common distinction is between *global* and *local* methods for sensitivity analysis (Morio 2011, Borgonovo & Plischke 2016). Global methods are based on quantifying the overall uncertainty around the output, and measuring how much of that uncertainty is due to particular inputs. Thus VoI can be characterised as a global method. Borgonovo et al. (2016) reviewed a range of alternative global sensitivity measures, and established mathematical conditions for a measure to be interpreted in terms of information value. Local methods describe the variations in the model output that would result from small perturbations in the model inputs around a base case (e.g. through calculating partial derivatives of the output with respect to the inputs). Thus they have been favoured for uncertainties that are hard to quantify, e.g. about different plausible prior distributions in Bayesian models (Roos et al. 2015). VoI, by contrast, requires defining parameters with probability distributions to represent uncertainty, which may be difficult if the evidence about those parameters is weak (O'Hagan & Oakley 2004).

4.1. Theory of VoI for point estimation: single parameters

To use Value of Information methods for sensitivity analysis in models used to estimate a quantity of interest α, estimation should be expressed as a decision problem (Bernardo & Smith 1994). Suppose we want to choose a point estimate $\hat{\alpha}$ for a scalar quantity. The set of alternative "actions" d consists of the allowable values for the parameter. The loss function typically describes a larger loss for estimates $d = \hat{\alpha}$ that are further from the true α. A natural choice is the squared error function. Then, since the loss $L(d, \boldsymbol{\theta})$ depends purely on $\alpha = g(\boldsymbol{\theta})$ and its estimate, we can write $L(d, \boldsymbol{\theta}) = L(\hat{\alpha}, \alpha) = (\hat{\alpha} - \alpha)^2$. The point estimate that minimises the expected loss can easily be shown, by differentiation, to be the mean of the current belief distribution $E_\alpha(\alpha)$. The expected loss given this decision is the variance of this distribution.

Hence the expected value of new information is then simply the expected *reduction in*

variance given the new information, as observed by Oakley & O'Hagan (2004). The EVPI, or total decision uncertainty, is equal to the current variance. The EVPPI for learning some parameter or parameters $\boldsymbol{\phi}$ (which could be any subset or function of $\boldsymbol{\theta}$) is

$$EVPPI_{\boldsymbol{\phi}} = \text{var}(\alpha) - \text{var}_{\alpha|\boldsymbol{\phi}}(\alpha) \qquad 6.$$

serves as a measure of sensitivity of the output α to the uncertain input(s) $\boldsymbol{\phi}$. This measure has also been called the *main effect* index of $\boldsymbol{\phi}$ (Oakley & O'Hagan 2004) or the *Sobol* index (after Sobol 1993), and is widely used without any interpretation in terms of decision theory or value of information (Saltelli et al. 2000).

This measure of information value, as reduction in variance of a belief distribution, extends immediately to describing the expected value of observable data $\mathbf{y}$ that reduces (rather than eliminates) uncertainty about the model parameter(s) $\boldsymbol{\phi}$. This is the expected value of sample information (3), which in this case is the expected reduction in variance,

$$EVSI_{\mathbf{y}} = \text{var}(\alpha) - \text{var}_{\alpha|\mathbf{y}}(\alpha) \qquad 7.$$

This extension is not generally discussed in literature on sensitivity analysis, though it is a familiar approach to *design of experiments*, as discussed in the following sections. The VoI perspective helps to give a unified framework for understanding similar concepts encountered in different fields of research. For example, Jackson et al. (2019) showed how a computational method developed for VoI in health economic decision problems can be generalised to a larger class of problems including the variance-based VoI formulae (6) and (7). This method, based on regression, will be described in more detail in Section 5.2.

4.2. Point estimates of multiple parameters

The idea of valuing information about a single, scalar parameter α by reductions in its variance can be generalised to vectors $\boldsymbol{\alpha}$ of multiple parameters. These ideas are central to the field of Bayesian design of experiments — for a review see, e.g. Chaloner & Verdinelli (1995). For example, the squared error loss generalises to a "quadratic" loss

$$L(\hat{\boldsymbol{\alpha}}, \boldsymbol{\alpha}) = (\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha})^T H (\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha})$$

where H is a symmetric, positive-definite matrix. If $H = \mathbf{c}\mathbf{c}^T$, for some $\mathbf{c}$, then the loss is a weighted sum of the parameters $\mathbf{c}^T\boldsymbol{\alpha}$. If all parameters are weighted equally, then the goal is to minimise an expected loss defined as the sum of all elements of the covariance matrix. Additionally if the parameters are independent, then the loss is the sum of their variances. Designs that minimise losses of this form are known as *A-optimal* designs. Here the same *absolute* reductions in variance are valued equally. For example, a decision maker would pay the same amount for information that reduces the variance from 200 to 100, as for information that reduces the variance from 2000 to 1900. These can be contrasted with *D-optimal* designs, which minimise an expected loss defined as the *determinant* of the covariance matrix. In those, the same *relative* reductions in variance are then valued equally, e.g. a reduction from 200 to 100 is valued the same as a reduction from 2000 to 1000.

Either way, the optimal design (disregarding study costs) is one that gives the highest expected value of sample information, for a "decision" problem defined as parameter estimation, and a specific choice of loss function. Accounting for costs would require the

decision-maker to specify the amount they are willing to pay for a specific variance reduction (see Jackson et al. (2019) for an example).

4.3. Estimating posterior distributions

Instead of simply point estimation, we could consider estimation of the *entire uncertainty distribution* of some quantity of interest $\boldsymbol{\alpha}$ as a decision problem. We could then use VoI to predict not only how well an experiment would estimate a quantity, but also how well we could characterise the *remaining uncertainty* (hence the value of even more information) after conducting the experiment.

Loss functions that describe how well a given probability distribution q describes a random variable $\boldsymbol{\alpha}$ are known as *scoring rules*. A *log scoring rule* is defined by the logarithm of the probability that q assigns to particular values of $\boldsymbol{\alpha}$. This forms the basis of a standard Bayesian decision-theoretic approach to design of experiments (following Lindley (1956), Bernardo & Smith (1994)), where gains in information are formulated as loss reductions or utility gains. In this framework, the expected gain in knowledge about $\boldsymbol{\alpha}$ given by future data $\mathbf{y}$, or the expected value of sample information for $\mathbf{y}$, is

$$E_{\boldsymbol{\alpha}}[-\log\{p(\boldsymbol{\alpha})\}] - E_{\mathbf{y}}E_{\boldsymbol{\alpha}|\mathbf{y}}[-\log\{p(\boldsymbol{\alpha}|\mathbf{y})\}]$$

the expected loss under current information, minus the expected loss after learning new data $\mathbf{y}$. This can be expressed in various equivalent ways, e.g. in terms of "mutual information" or Kullback-Leibler divergences, see, e.g. Ryan et al. (2016).

The expected value of partial perfect information for a parameter ϕ would have the same form, but with ϕ instead of $\mathbf{y}$. The expected value of perfect information (the expected loss under the current belief distribution) would be $E_{\boldsymbol{\alpha}}(-\log(p(\boldsymbol{\alpha})))$, known as the *entropy* of the distribution p. The entropy describes the current level of uncertainty in a distribution, but in a different way from the variance, describing how *concentrated* the distribution is. For example, a distribution with two concentrated peaks that are far apart from each other will have low entropy but high variance, while a distribution with just one concentrated peak will have low entropy and low variance.

For estimating the parameters of a normal linear model, maximising the expected value of sample information in this form is equivalent to minimising $\det(\mathrm{cov}(\boldsymbol{\alpha}))$, giving a D-optimal design (Chaloner & Verdinelli 1995). In other situations, computing the nested expectations in these formulae is challenging, as discussed by Ryan et al. (2016), particularly when Markov Chain Monte Carlo (MCMC) is required to calculate the posterior $p(\boldsymbol{\alpha}|\mathbf{y})$. Thus this approach to experimental design, as a problem of estimating the *whole distribution* of $\boldsymbol{\alpha}$, has not been as common in practice, compared to the approach based on point estimation of $\boldsymbol{\alpha}$ and variance-based losses.

This article just gives a brief illustration of the link between the ideas of VoI and Bayesian design of experiments. There is a much wider literature on Bayesian design, see, e.g. Woods et al. (2017) for a recent review and some computational developments, Drovandi et al. (2017) for a discussion of sampling design issues in the context of very large datasets, and Laber et al. (2018) for an example of optimal design in dynamic sampling situations where a large number of alternative designs must be considered.

4.4. Examples of VoI for estimation models in health

4.4.1. VoI in mathematical models: health impacts of transport. A health-related example of deterministic mathematical models of the form $\alpha = g(\boldsymbol{\theta})$ is given by the models used to determine the health impacts of scenarios where the usage of different forms of transport changes for a population (Mueller et al. 2015). Health impacts of changes in transport behaviour may come from physical activity benefits of cycling and walking, changes in road traffic injuries or changes in air and noise pollution. A policy-maker may want to know, for example, the health impacts of car use reducing by a half, before considering any policy to achieve that change. The model parameters $\boldsymbol{\theta}$ include, e.g. parameters describing pollution exposure from cars and physical activity benefits of cycling, the relative risks of various chronic diseases associated with physical activity and air pollution, and measures of health-related quality of life for different diseases used to obtain an overall health outcome α (e.g. disability-adjusted life years).

An application of EVPPI in a model of this kind was given by Johnson et al. (2019), showing how better evidence on particular pollution-related parameters would give more precise estimates of the number of deaths $\mathbf{y}$ prevented per year in an scenario of increased active transport in São Paulo (de Sá et al. 2017).

4.4.2. VoI in multiparameter evidence synthesis: HIV prevalence estimation. Jackson et al. (2019) used Value of Information methods in a model for estimating the prevalence of HIV infection in London among men who have sex with men. Since the quantities of interest are not directly observable, surveys, register datasets and expert judgements were combined in a Bayesian multiparameter evidence synthesis model, fitted by MCMC methods.

An EVPPI analysis showed which uncertain parameters contributed most to the uncertainty in the different prevalence estimates required. This can be illustrated with "heat maps", as in Figure 1 (left). The columns indicate nine different outputs produced from the model: the proportion of people with diagnosed infection (indicated $(\pi\delta)$), the proportion with undiagnosed infection $\overline{(\pi\delta)}$, and the corresponding numbers of people with infection μ. The subscripts D and U refer to diagnosed and undiagnosed infection, and the subscripts G and N indicate two different risk groups, those who have or haven't attended a GUM (sexual health) clinic in the past year, respectively. μ with no subscript indicates the total number infected. The lighter squares indicate which of the uncertain input parameters (rows) contribute most to the uncertainty about eight alternative output quantities of interest (columns). For example, uncertainty about the undiagnosed number μ_{UN} and prevalence $\overline{(\pi\delta)}_N$ among non-GUM attenders is driven most by the parameter $or^{(GM)}$, the relative odds of having HIV between non-GUM and GUM attenders.

An EVSI analysis (Figure 1, right) then shows the expected value of further survey data of a given sample size from the same source originally used to estimate $or^{(GM)}$ (labelled "GMSHS data"). Modest reductions in variance of μ_U (the overall undiagnosed number) would be expected from obtaining 100-1000 extra observations from this source. Another survey data source (labelled "GUM Anon data") is judged to lead to much smaller reductions in the variance of μ_U for the same sample size — while Figure (left) showed that the parameter estimated from the GUM Anon data $\pi^{(GA)}$ explained most of the uncertainty about the undiagnosed prevalence in GUM attenders $\overline{(\pi\delta)}_G$, the EVSI analysis shows that extra data from GUM Anon would have little benefit in improving the precision of μ_U.

Finally, the ENBS can be calculated to account for the costs of data collection, and hence to determine an optimal sample size. This requires the decision maker to determine

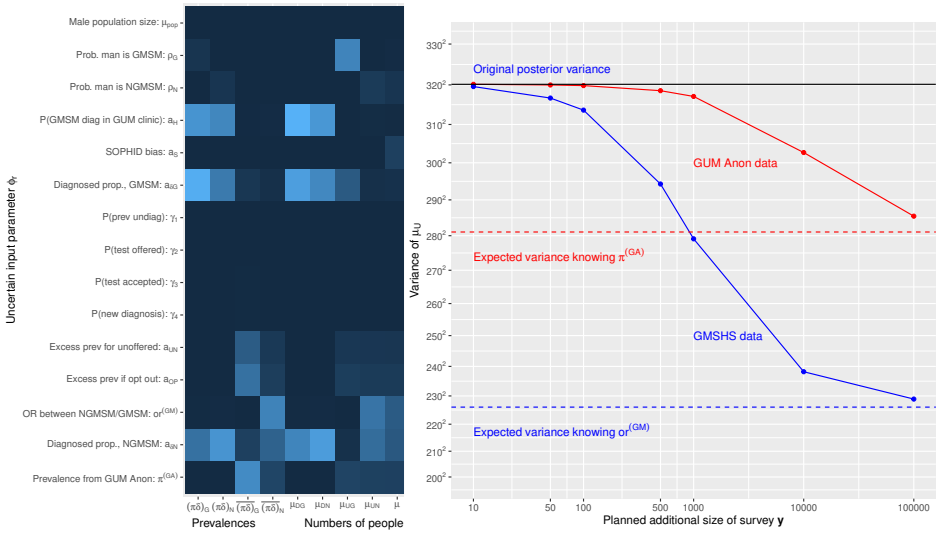

Figure 1

Illustration of an VoI analysis for an estimation model, from Jackson et al. (2019). The left panel illustrates the EVPPI. Rows indicate uncertain model input parameters ϕ. Columns indicate different "output" quantities α that the model was designed to estimate; the prevalence, and number of people with, HIV infection in London in various risk groups, considering diagnosed and undiagnosed infections separately. Lighter coloured squares indicate higher $EVPPI_{\phi_r}/\text{var}(\alpha_s)$, the expected proportion of variance of the output α_s that would be reduced by knowing the exact value of the input ϕ_r. The right panel indicates EVSI in terms of expected reduction in variance of a specific model output μ_U after collecting additional data of given sample sizes from two alternative sources ("GUM Anon" and "GMSHS"), which is bounded by the EVPPI for the parameter informed by that data.

the amount that they are willing to pay for a specific improvement in precision of the estimate, as described in Jackson et al. (2019), which may be difficult to elicit in practice. However, a more formal cost-effectiveness analysis could be performed if we were using these prevalence estimates as part of evaluating a health policy, e.g. an intervention to increase the proportion of people with HIV infection who are diagnosed. More precise estimates of prevalence will lead to better estimates of the benefit of such a policy, hence a better-informed decision for whether to implement it. The net health benefit of the policy, the value of further information to improve estimates of prevalence, and the optimal study design, could be determined using the methods described in Section 3. In that framework, benefits of the policy and of further information can both be measured on the same scale as health. Hence the ENBS calculation is easier, since the decision-maker specifies their willingness to pay for a health benefit (typically one QALY), which is better-understood than the willingness to pay to reduce the variance of an estimate.

4.5. Open issues with VoI in models used for estimation

Open issues in the general field of sensitivity analysis for complex models are discussed elsewhere (e.g. Borgonovo & Plischke 2016, Razavi et al. 2021). As we have discussed, VoI methods require the extent of uncertainty to be described with a specific probability distribution, which may be difficult if the information is weak. An advantage of the VoI

approach to sensitivity analysis is the way that it extends naturally to designing further data collection.

A general challenge to VoI arises when the model is used to estimate more than one quantity, and a proposed study might be informative about multiple quantities. This is a common situation in the transport and health models discussed above, where the model can describe health impacts of different kinds to different populations. To prioritise further data collection when the data informs multiple model outcomes, we would have to either calculate a VoI estimate for each of the outcomes and qualitatively judge their relative importance, or formulate an appropriate loss function to obtain an overall measure of value. More experience is needed to clarify how VoI can be applied in particular practical contexts here.

Mathematical models can take many different forms. Often they are computationally expensive, which can make any form of sensitivity analysis difficult, as discussed further in Section 5.2.3. Multiparameter evidence syntheses are harder in general compared to mathematical models that just represent a mechanistic process, since they add an extra step of Bayesian inference of parameters from data assumed to arise from part of the process (often described as model "calibration"). As well as computational difficulties, a *statistical* challenge for VoI occurs in "microsimulation" models, where the outcomes of interest are estimated by simulating and summarising a large synthetic population. Characteristics of individuals in the synthetic population are often determined by sampling with replacement from an observed individual-level dataset. See, e.g. Mytton et al. (2018) for an example from health policy modelling. In such cases, uncertainty about the *average* characteristics of those individuals (note the distinction from individual-level variability) is not being represented by a distribution on a parameter, but implicitly through the finite size of the observed dataset. To enable VoI methods, which require representing uncertainty through parameters, it may be necessary to construct a parametric model, fitted to the observed data, from which synthetic individuals could be sampled.

5. Computation of Value of Information measures

As is typical for a generic Bayesian statistical method, a major challenge to applying it routinely is the computation. This section reviews the available methods and challenges with computing each of the measures of value of information. Unless specified otherwise, the methods discussed here were developed for health economic decision problems.

Suppose we can obtain a sample of values $\boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(R)}$ from the current uncertainty distribution of the parameters, either randomly, or through quasi-Monte Carlo methods (Caflisch 1998), that generate non-random samples to cover the sample space. The EVPI (1) can then be computed straightforwardly by Monte Carlo integration, as an empirical average "opportunity loss" over the sample, $\frac{1}{R}\sum_{r=1}^{R}\{L(d^*, \boldsymbol{\theta}^{(r)}) - L(d^*_{\boldsymbol{\theta}}, \boldsymbol{\theta}^{(r)})\}$, see, e.g. Baio (2018) for an illustration. This is routinely applied in health economic decision problems (Briggs et al. 2006). The only concern is the extent of Monte Carlo error from the choice of the number of simulations R. This may not be negligible in situations where the distribution of $\boldsymbol{\theta}$ is awkwardly shaped or otherwise expensive to characterise by sampling, e.g. if using MCMC, so some sensitivity analysis to the choice of R is wise. Quasi-Monte Carlo methods may also help to reduce this error (Fang et al. 2021).

5.1. Expected value of partial perfect information

A traditional "brute force" approach to calculating the EVPPI in finite decision problems (equations 2,4) involves a nested Monte Carlo integration implemented as a "loop within a loop". A sample of values $\boldsymbol{\phi}^{(1)}, \ldots, \boldsymbol{\phi}^{(R)}$ is drawn for the parameter(s) of interest $\boldsymbol{\phi}$ from its uncertainty distribution. Each sample $\boldsymbol{\phi}^{(r)}$ mimics a situation where we have "partial perfect information": knowledge that $\boldsymbol{\phi} = \boldsymbol{\phi}^{(r)}$. Then for each r, we produce an estimate $e(\boldsymbol{\phi}^{(r)})$, of $E_{\boldsymbol{\theta}|\boldsymbol{\phi}^{(r)}}(L(d^*_{\boldsymbol{\phi}^{(r)}}, \boldsymbol{\theta}))$, the expectation of the loss of the decision taken given the information that $\boldsymbol{\phi} = \boldsymbol{\phi}^{(r)}$, with respect to the conditional uncertainty distribution of the remaining parameters $\boldsymbol{\theta}|\boldsymbol{\phi}^{(r)}$. This is done is obtained by drawing a sample, $\boldsymbol{\theta}^{(r,s)} : s = 1, \ldots, S$, from the conditional distribution $\boldsymbol{\theta}|\boldsymbol{\phi}^{(r)}$, and taking the empirical average $e(\boldsymbol{\phi}^{(r)}) = \frac{1}{S}\sum_{s=1}^{S} L(d^*_{\boldsymbol{\phi}^{(r)}}, \boldsymbol{\theta}^{(r,s)})$ (the "inner" loop). The second term of (2) can then be computed as $\frac{1}{R}\sum_{r=1}^{R} e(\boldsymbol{\phi}^{(r)})$ an average over the r (the "outer" loop), reflecting that we do not actually know $\boldsymbol{\phi}$ when we compute the EVPPI. The first term of (2) can be calculated with a single loop, in a similar way to the EVPI.

While the biases and error variances associated with such Monte Carlo approximations can be estimated (Oakley et al. 2010), the brute force method is impractical for all but the cheapest of problems, requiring the loss $L(\cdot)$ to be evaluated for tens of millions of samples. Therefore many more efficient approaches have been investigated, which mostly involve calculating or approximating the inner expectation, so that the computation reduces to a single Monte Carlo loop. This can be done algebraically, if the form of the function $L(\cdot)$ is known and satisfies certain linearity assumptions, and/or if parameters are jointly distributed in a particular way (Coyle & Oakley 2008, Madan et al. 2014). This approach was used to calculate EVPPI in the surgical dressing study (Reeves et al. 2019) discussed in Section 3, which involved a relatively simple net benefit function.

However, the algebraic approaches require a different form of computation to be derived for each different decision problem. This has motivated methods that can instead be implemented in general-purpose software. General-purpose methods that are only applicable when $\boldsymbol{\phi}$ represents one scalar parameter were developed by Strong & Oakley (2013) and Sadatsafavi et al. (2013).

5.1.1. Nonparametric regression methods. A more general approach to EVPPI computation is based on *nonparametric regression*. This was devised by Strong et al. (2014) to estimate (4) for health economic models, it is applicable when $\boldsymbol{\phi}$ includes multiple parameters, and has been implemented in an R package and web application (`https://savi.shef.ac.uk/SAVI`). This involves expressing the true net benefit as a function of $\boldsymbol{\phi}$:

$$-L(d, \boldsymbol{\theta}) = NB_d(\boldsymbol{\theta}) = E_{\boldsymbol{\theta}|\boldsymbol{\phi}}(NB_d(\boldsymbol{\theta})) + \epsilon = g_d(\boldsymbol{\phi}) + \epsilon \qquad 8.$$

for each $d = 1, \ldots, D$, where ϵ is an error term with mean zero. Then $g_d(\boldsymbol{\phi})$ is estimated by a flexible regression of NB_d on $\boldsymbol{\phi}$, fitted to a "dataset" defined by a random sample of $(NB_d^{(r)}(\boldsymbol{\theta}), \boldsymbol{\phi}^{(r)}) : r = 1, \ldots, R$ from their current belief distribution. If the parameters of interest $\boldsymbol{\phi}$ comprise p quantities that we want to learn the joint value of perfect information on, the regression will have p predictors. This produces a fitted value $\hat{g}_d(\boldsymbol{\phi}^{(r)})$ for each r, which allows the second term in (4) to be estimated by a Monte Carlo mean over the random sample,

$$E_{\boldsymbol{\phi}} \max_d \{E_{\boldsymbol{\theta}|\boldsymbol{\phi}}(NB_d(\boldsymbol{\theta}))\} \approx \frac{1}{R}\sum_{r=1}^{R} \max_d \{\hat{g}_d(\boldsymbol{\phi}^{(r)})\}. \qquad 9.$$

The nonparametric regression method was shown by Heath et al. (2017) to be easier to implement in single-parameter situations than the methods of Strong & Oakley (2013) and Sadatsafavi et al. (2013). The main challenge for the regression method is how to specify the regression function g_d, since the functional form will not be known in general. Most implementations have been based on either generalized additive models (GAMs) or Gaussian processes (GPs). Strong et al. (2014) suggested GAMs when the number of parameters p was less than around five, and GPs otherwise.

A wide range of methods are available to construct GAMs, typically based on splines (Wood 2017), though their relative merits for EVPPI computation have not been systematically compared. Jackson et al. (2019) found multivariate adaptive regression splines (essentially based on linear splines) as implemented in the `earth` R package (Friedman 1991, Milborrow 2011), to be more efficient than cubic splines in one particular application. A disadvantage of the additive regression form of GAMs is that the potential number of models that can be constructed with p regressors (including multi-way interactions) increases sharply with p, with the largest possible model typically becoming prohibitive as p increases beyond 5. The optimal level of complexity for the regression is therefore uncertain. This is a problem of statistical identifiability as well as computational speed — for more complex models to become identifiable, larger simulated "datasets" (i.e. with bigger R) must be generated.

Gaussian process models construct the regression function $g(x)$ so that for any choice of $x_1, x_2, \ldots, x_n$, $(g(x_1), \ldots, g(x_n))$ has a multivariate normal distribution, with a mean and covariance typically constructed to express higher correlations when x_i is closer to x_j for each i, j. Fitting a GP model typically involves choosing or estimating a parameter that determines the smoothness of $g(\cdot)$. Since fitting GP models involves linear algebra on $R \times R$ matrices, these models scale badly when a large number of samples R is required to characterise the distribution of parameters $\boldsymbol{\theta}$, though there are efficient approximations based on inverting smaller matrices (Quinonero-Candela & Rasmussen 2005). In our experience, standard methods (Strong et al. 2014) are only practicable for R up to around 10000. Heath et al. (2016) developed a more efficient Gaussian process method for EVPPI calculation, based on spatial modelling principles, which projects the p predictors into a lower-dimensional space (see Baio et al. (2017) for more details).

In all regression models, the standard error from the regression coefficients can be propagated to give a standard error for the estimate of the EVPPI, along with any Monte Carlo error from the mean (9), though note that this does not acknowledge uncertainty about the choice of regression model form.

5.1.2. Computing EVPPI for reductions in variance in estimation problems. The regression idea can also be used to calculate the EVPPI (Equation 6) for mathematical models to estimate a quantity of interest α (Stanfill et al. 2015). This applies both to models where α is a known deterministic function of inputs $\boldsymbol{\theta}$, and multiparameter evidence synthesis models where we estimate a joint posterior distribution of $\alpha, \boldsymbol{\theta}$, as long as a sample from this distribution can be obtained (Jackson et al. 2019). Note that an alternative common method (Sobol 1993) for computing measure (6) requires the model to be a known deterministic function, thus does not apply to multiparameter evidence synthesis models.

To estimate the EVPPI for a particular parameter or parameters $\boldsymbol{\phi}$, we simply fit a single regression model $\alpha = g(\boldsymbol{\phi}) + \epsilon$ to this sample $(\alpha^{(r)}, \boldsymbol{\phi}^{(r)}) : r = 1, \ldots, R$, just like Equation (8). The fitted value $\hat{g}(\boldsymbol{\phi}^{(r)})$, at a value $\boldsymbol{\phi}^{(r)}$, estimates $E(\alpha|\boldsymbol{\phi} = \boldsymbol{\phi}^{(r)})$, the

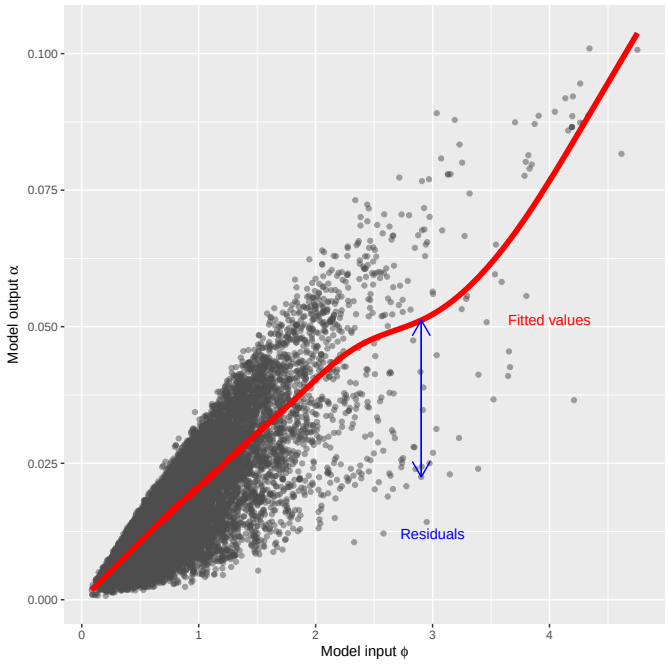

Figure 2

Illustration of using regression to estimate EVPPI of a scalar model input $\boldsymbol{\phi}$, as the expected reduction in variance of a model output α after learning $\boldsymbol{\phi}$, using a sample from the joint distribution of α and $\boldsymbol{\phi}$. The fitted value at $\boldsymbol{\phi}^{(r)}$ is $E(\alpha|\boldsymbol{\phi} = \boldsymbol{\phi}^{(r)})$. The variance of the fitted values is the EVPPI, and the variance of the residuals is the variance remaining after learning α, $\text{var}(E(\alpha|\boldsymbol{\phi} = \boldsymbol{\phi}^{(r)}))$

expected outcome if the uncertain parameter $\boldsymbol{\phi}$ were known to equal $\boldsymbol{\phi}^{(r)}$. Conversely, the residuals $\alpha^{(r)} - \hat{g}(\boldsymbol{\phi}^{(r)})$ represent the variability in the outcome that is not explained by π. In other words, the residual variance σ quantifies the uncertainty about α that would remain if we were to learn the value of π. Therefore we can estimate the EVPPI

$$\text{var}(\alpha) - E_{\boldsymbol{\phi}}\left[\text{var}_{\alpha|\boldsymbol{\phi}}(\alpha)\right] = \text{var}_{\boldsymbol{\phi}}\left[E_{\alpha|\boldsymbol{\phi}}(\alpha|\boldsymbol{\phi})\right]$$

as the empirical variance of the fitted values $\hat{g}(\boldsymbol{\phi}^{(r)})$, roughly interpreted as the variance "explained by" $\boldsymbol{\phi}$. This is illustrated in Figure 5.1.2. Equivalently we could calculate the variance of the $\alpha^{(r)}$ minus the residual variance, an estimate of the reduction in variance on learning $\boldsymbol{\phi}$.

Just as for decision problems, this method is simple and generally applicable, though there is the same risk of sensitivity to regression model choice.

5.1.3. Summary of challenges for EVPPI computation. In many situations, the currently-available methods for EVPPI computation are accurate enough to make decisions of how to prioritise further research. Research can be ruled out for a quantity ϕ if the EVPPI is low. If the EVPPI is high, we may want to go further and calculate the EVSI. Challenges remain for larger problems, either where we require a joint EVPPI for a large number p of parameters (more than around 15, in our experience) or where large Monte Carlo sample sizes R are

required. In health economic evaluations, p might be large if we are considering a proposed trial from which we estimate not only treatment effects on a primary outcome, but also other parameters required for decision modelling, such as costs and health-related quality of life for different disease states, or probabilities of transition between these states. Large sample sizes may be required if using MCMC methods, e.g. if current evidence about treatment effects comes from Bayesian evidence synthesis models such as network meta-analysis. More advanced "multilevel" Monte Carlo methods may help in these situations (Fang et al. 2021, Giles & Goda 2019).

Other challenges arise where the loss or net benefit function is computationally expensive. These include *microsimulation* models, where policies or scenarios are compared by generating outcomes from a large sample of synthetic individuals and summarising the sample. There has been limited experience of VoI methods for these, though *emulation* or *meta-modelling* techniques have been suggested as a general approach for expensive models (Oakley & O'Hagan 2004).

EVPPI computation: summary of methods and challenges

1. Brute force Monte Carlo methods can give answers to arbitrary accuracy in theory, but are typically too expensive.
2. Algebraic simplifications may work, but require model-specific derivation and programming.
3. Nonparametric regression is easy to implement, efficient and widely applicable, though there may be uncertainty about the choice of nonparametric model.
4. Challenges remain for many parameters, large sample sizes and expensive models.

5.2. Expected value of sample information

Computing the EVSI (equations 3,5) is conceptually harder than computing the EVPPI, because it requires an extra step of Bayesian inference: considering how knowledge of θ will be updated given data $\mathbf{y}$ that are as yet unknown. A basic requirement is to specify a data-generating model for $\mathbf{y}$ under the proposed study design. Methods for calculating EVSI then generally involve simulating from the predictive distribution for $\mathbf{y}$ given current information on θ — achieved by simulating θ, followed by simulating $\mathbf{y}|\theta$.

As for EVPPI, brute-force methods have been used (Ades et al. 2004), where alternative parameters $\boldsymbol{\theta}^{(r)}$, hence alternative datasets $\mathbf{y}^{(r)}|\boldsymbol{\theta}^{(r)}$, are simulated in an "outer" loop, each $\mathbf{y}^{(r)}$ mimicking a situation where we have new information. The posterior distribution given this new information $\boldsymbol{\theta}|\mathbf{y}^{(r)}$ is then determined, and in an "inner loop", we simulate from this posterior to determine an updated optimal decision and expected loss $E_{\boldsymbol{\theta}|\mathbf{y}^{(r)}}(L(d^*_{\mathbf{y}^{(r)}}, \boldsymbol{\theta}))$. The computational burden is similar to that of the brute-force EVPPI method, as long as we can sample instantly from the the posterior distribution of the expected loss given a simulated dataset. As this may require intensive methods, such as MCMC, it may be more expensive than calculating EVPPI.

As for EVPPI, more advanced "multilevel" Monte Carlo methods have been proposed (Hironaka et al. 2020). In specific modelling situations, as for EVPPI, the brute force formula may also reduce to a version that is simpler to compute (Ades et al. 2004). Likewise, in certain situations, methods for approximating Bayesian inference, e.g. Laplace approxi-

mations, might be used to calculate EVSI (Brennan & Kharroubi 2007a,b). For example, a normal approximation to the posterior was used to derive an EVSI calculation in the surgical dressing case study (Reeves et al. 2019). However this approach requires the practitioner to derive and program a different analytic formula for each situation. Therefore a variety of more efficient methods have been proposed in recent years, intended to either be computationally cheaper, or more amenable to general-purpose software implementations.

5.2.1. Nonparametric regression. Strong et al. (2015) adapted a simple extension of the regression-based EVPPI computation method to compute EVSI. The basic assumption is that all information provided by the data $\mathbf{y}$ about the parameters θ can be expressed as an easily-computable summary statistic $T(\mathbf{y})$. Then Equation 8 becomes

$$NB_d(\boldsymbol{\theta}) = g_d(T(\mathbf{y})) + \epsilon$$

and $g_d(\cdot)$ is estimated by a nonparametric regression fitted to a random sample with outcome $NB_d^{(r)}$ and predictor $T(\mathbf{y}^{(r)})$, $r = 1, \ldots, R$. If p different summary statistics are required to summarise the relevant information in the data, then there will be p predictors in the regression. The EVSI is then calculated using Equation 9 with $\boldsymbol{\phi}^{(r)}$ replaced by $T(\mathbf{y}^{(r)})$.

This method involves only a single level of Monte Carlo simulation, so is vastly more efficient than the brute force method, as long as the function $T(\cdot)$ can be evaluated quickly. It is amenable to implementation in general-purpose software, and widely applicable in healthcare evaluation, where information is commonly obtained from a randomised trial that is designed to be simple to summarise. For example, there might be two treatment groups of size N_1, N_2, where the data consist of the number y_1, y_2 that experience some binary outcome. The limitations of this method are similar to those for the analogous method for EVPPI, e.g. where the proposed study gives information about a large number p of parameters of interest, thus fitting a sufficiently-flexible regression may be expensive. Estimates might also be sensitive to different choices of $T(\cdot)$, just as the choice of regression function might be influential.

A study designed to give the desired information, e.g. a randomised trial, will typically be designed in such a way that the quantity of interest $T(\cdot)$ can be computed easily. This may not be the case if the required analysis is complicated, for example, with observational data where there are biases expected (e.g. selection, confounding or informative missingness) that must be adjusted for by analysis, or if the required data have a complex (e.g. hierarchical) structure.

5.2.2. Moment-based and other general-purpose methods. Another approach to EVSI calculation is based on assuming that information gains are completely specified by reductions in variance of some quantity of interest. If this variance reduction can be estimated, then this allows EVPPI calculation procedures to be modified easily to estimate the EVSI.

Heath et al. (2019) used this idea, described as "moment matching". A small training sample is used to estimate the expected posterior variance (given sample information) of the net benefit, ${\sigma^2}_{new} = E_{\mathbf{y}}\{\mathrm{var}_{\boldsymbol{\theta}|\mathbf{y}}(NB_d(\boldsymbol{\theta}))\}$. This is compared to the variance $\sigma^2 = \mathrm{var}_{\boldsymbol{\theta}}(NB_d(\boldsymbol{\theta}))$ of the net benefit under current information. The proportion of uncertainty ${\sigma^2}_{new}/\sigma^2$ expected to remain after the proposed study is then used to rescale the estimates of the conditional expected net benefit from the EVPPI calculation, e.g. the fitted values from the regression-based method (9), to estimate the EVSI. A similar approach was developed by Jalal & Alarid-Escudero (2018), based on determining reductions in variance

of parameters $\boldsymbol{\theta}$, rather than the net benefit. The most difficult part of these approaches is the Bayesian inference step. Jalal & Alarid-Escudero (2018) observed that the prior to posterior information gain can be calculated easily in cases where the prior and posterior are from the same conjugate family. Otherwise, MCMC methods may be required. As a reasonable accuracy may be achieved with a training sample of 50 or less (Heath et al. 2019) these approaches are less expensive than the brute force method, but more expensive than the "summary statistic" regression method of Strong et al. (2015). However they are generalisable to situations where there is no summary statistic, and negligible extra computation is required if the EVSI is required for multiple designs that differ only in their sample size.

A further method was developed by Menzies (2016) based on importance sampling, which involves computing the likelihood for a matrix of combinations of simulated datasets and parameter values. This is conceptually simple, but more expensive than the regression-based method. Heath et al. (2020) compared the regression, variance-based and importance sampling methods in three realistic health economic models, concluding that no method was consistently superior in terms of efficiency and accuracy, and their relative advantages depended on the modelling context. Kunst et al. (2020) gave a practical tutorial in these methods, explaining the skills that are required to implement them.

5.2.3. Computation for EVSI in estimation problems and optimal design problems. The methods previously discussed in Section 5.2 were all developed for health economic decision models used to choose between a set of policies. In estimation problems, since the EVSI is equivalent to the expected reduction in loss from an experiment, methods for calculating it are covered in the literature on Bayesian experimental design. See Ryan et al. (2016) for a recent review.

However, Bayesian design traditionally deals with situations where there is is a single source of experimental data, combined with an explicit prior. In contrast, the models used in the VoI literature are typically based on synthesis of multiple sources of data (due to the need to consider all relevant data in evidence-based medicine) hence there are a large number of uncertain parameters to be considered when predicting gains from further data. For such models, an EVSI calculation would then normally follow an EVPPI calculation where the most influential parameters are determined. It would then be convenient to use a similar method to implement both parts of the VoI analysis. This could be done with the regression method described in 5.2.1, which generalises for computing the EVSI in estimation problems, just as the analogous method for computing EVPPI generalises (Section 5.1.2). Note also that the moment-based methods are explicitly designed to quantify expected reductions in variance, so could in principle be adapted for estimation problems.

An extra dimension of difficulty arises when the space of potential designs to be compared is large. In health, this might happen in certain adaptive clinical trial designs, where decisions on how to allocate participants to treatments are made at a sequence of time points, based on information accumulated up to that point. See, for example, Villar et al. (2015) for a review of Bayesian methods for optimal designs in such a context. See also Conti & Claxton (2009) for an example of a complex design space in health economics. In surveillance and control of infectious diseases, we might want to choose both times and spatial locations for collecting further data, see, e.g. Laber et al. (2018) for a recent example. A further difficulty common in infectious disease epidemiology is that the model representing transmission may be computationally expensive, particularly if based on simulating at an individual-level. Emulation approaches, as in e.g. Andrianakis et al. (2015), might be

extended to allow the value of further information to be determined.

Computation of the EVSI in evidence synthesis models: summary of methods and challenges

1. Brute force Monte Carlo methods are expensive.
2. EVSI computation can sometimes be simplified algebraically, but this requires model-specific derivation and programming.
3. Nonparametric regression is is easy and efficient, and applicable if we can describe the sample information as a summary statistic, though care is required with potential sensitivity to the regression form and summary function. Otherwise, methods based on variance reduction can be useful.
4. Challenges remain for optimal design with complex design spaces or with expensive models.

6. Discussion

Bayesian decision theory is an elegant framework for combining evidence in models to inform scientific investigations and policy-making. This article has reviewed how Value of Information methods are used in this framework: to find out where the weakest evidence in a decision model is, and what data should be collected to improve it.

While the idea of decision theory is elegant, a clear difficulty with using it is that it requires the problem and evidence to be expressed in a very formal, quantitative manner. As discussed in the context of health economics, models should include all relevant data and quantify all benefits and harms of any resulting decisions. *Multicriteria decision analysis* is an extensive field of research that deals with comparing multiple benefits or harms quantitatively, see, e.g. Ishizaka & Nemery (2013) for a review, and Nutt et al. (2010) for a well-known application to the harms of recreational drugs. Another health-related application of formal decision analysis is in balancing efficacy and safety in pharmaceutical research, see Hughes et al. (2016) for a review of methods.

Another general challenge is how to quantify uncertainty in models. A potential dilemma occurs where we want to use VoI methods to guide improvements in information that is too weak to be quantified confidently, but in order to use VoI, we need to specify full probability distributions for uncertain parameters. In such situations, we might use a range of reasonable "vague" uncertainty distributions, and hope that the estimates of VoI lead to the same research prioritisation decision, e.g. that further research is not worthwhile. This is a similar approach to using a range of priors for Bayesian inference, as suggested by, e.g. Spiegelhalter et al. (2004) in the context of clinical trials, where a conclusion from data might be strengthened if it doesn't change when combined with a "sceptical" prior. Structured expert elicitation (O'Hagan et al. 2006) may also help to determine a reasonable range of beliefs. The "Bayes linear" approach, in which prior information is expressed and updated through means, variances and covariances, rather than full probability distributions, may also be useful for dealing with subjective information, and has been used in experimental design (Jones et al. 2016). Note also that data gathering may be iterative or sequential – a small "pilot" experiment, as described in the health economic evaluation example (Reeves et al. 2019), can help to more confidently establish the value of stronger information.

Finally note that a purely Bayesian perspective has been taken throughout, due to the convenient way in which this approach deals with multiple sources of evidence and uncertainty. A Bayesian approach is not strictly necessary for decision theory, and some frequentist alternatives are described in, e.g. Parmigiani & Inoue (2009), and Berger (2013). The frequentist perspective for experimental design is well known, e.g. clinical trials are usually designed and analysed by considering error rates of hypothesis tests. This does not explicitly compare the expected benefits and harms of different decisions, while often involving uncertain quantities such as the minimum clinically important difference (Chuang-Stein et al. 2011). Manski (2019) advocated (frequentist) decision theory as an alternative to hypothesis testing in the analysis of clinical trial data. However, the connections of frequentist decision theory with value of information for evidence synthesis have not been explored. Clearer links between all of these areas would help to establish that much of statistical science is meant as a basis for decision making, and help to emphasise the importance of considering the value of further evidence.

A general purpose software package to implement the most widely-applicable of the computational methods for EVPPI and EVSI discussed in Section 5 is under development at `http://chjackson.github.com/voi`.

Value of Information: challenges and issues

- VoI requires formalising a decision problem and building a model to represent all relevant evidence.
- Benefits and harms may be challenging to model and quantify.
- Uncertainties about evidence in models can be challenging to represent as specific probability distributions.
- Computation can be difficult for expensive models and when searching over a large space of study designs.
- Cross-fertilising VoI with similar ideas from different fields, both Bayesian and frequentist, could help to strengthen the methodology, and to emphasise the importance of decision-making and evidence gathering in statistical science.

7. Acknowledgements

Thanks to all the members of ConVoI (Collaborative Network for Value of Information), and to Howard Thom for helpful suggestions. The first author was funded by the Medical Research Council, programme number MRC_MC_UU_00002/11.

LITERATURE CITED

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

Andrianakis I, Vernon IR, McCreesh N, McKinley TJ, Oakley JE, et al. 2015. Bayesian history matching of complex infectious disease models using emulation: a tutorial and a case study on HIV in Uganda. *PLoS Computational Biology* 11:e1003968

Arrow KJ, Lind RC. 1970. Uncertainty and the evaluation of public investment decisions. *The American Economic Review* 60:364–378

Baio G. 2018. Statistical modeling for health economic evaluations. *Annual Review of Statistics and its Application* 5:289–309

Baio G, Berardi A, Heath A. 2017. Bayesian Cost-Effectiveness Analysis with the R Package BCEA. Springer

Berger JO. 2013. Statistical decision theory and Bayesian analysis. Springer

Bernardo JM, Smith AFM. 1994. Bayesian theory. Chichester: Wiley

Borgonovo E, Hazen GB, Plischke E. 2016. A common rationale for global sensitivity measures and their estimation. *Risk Analysis* 36:1871–1895

Borgonovo E, Plischke E. 2016. Sensitivity analysis: a review of recent advances. *European Journal of Operational Research* 248:869–887

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

Brennan A, Kharroubi SA. 2007a. Efficient computation of partial expected value of sample information using Bayesian approximation. *Journal of Health Economics* 26:122–148

Brennan A, Kharroubi SA. 2007b. Expected value of sample information for Weibull survival data. *Health Economics* 16:1205–1225

Briggs A, Sculpher M, Claxton K. 2006. Decision modelling for health economic evaluation. Handbooks in Health Economic Evaluation. Oxford: Oxford University Press

Briggs AH, Weinstein MC, Fenwick EAL, Karnon J, Sculpher MJ, et al. 2012. Model parameter estimation and uncertainty: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6. *Value in Health* 15:835–842

Caflisch RE. 1998. Monte Carlo and quasi-Monte Carlo methods. *Acta Numerica* 1998:1–49

Caro JJ, Briggs AH, Siebert U, Kuntz KM. 2012. Modeling good research practices—overview: a report of the ispor-smdm modeling good research practices task force–1. *Medical Decision Making* 32:667–677

Chaloner K, Verdinelli I. 1995. Bayesian experimental design: A review. *Statistical Science* :273–304

Chuang-Stein C, Kirby S, Hirsch I, Atkinson G. 2011. The role of the minimum clinically important difference and its impact on designing a trial. *Pharmaceutical Statistics* 10:250–256

Claxton K. 1999. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *Journal of Health Economics* 18:341–364

Claxton K, Neumann PJ, Araki A, Weinstein MC. 2001. Bayesian value-of-information analysis: an application to a policy model of Alzheimer's disease. *International Journal of Technology Assessment in Health Care* :38–55

Claxton KP, Sculpher MJ. 2006. Using value of information analysis to prioritise health research. *Pharmacoeconomics* 24:1055–1068

Conti S, Claxton K. 2009. Dimensions of design space: a decision-theoretic approach to optimal research design. *Medical Decision Making* 29:643–660

Coyle D, Oakley J. 2008. Estimating the expected value of partial perfect information: a review of methods. *European Journal of Health Economics* 9:251–259

Davis DR, Kisiel CC, Duckstein L. 1972. Bayesian decision theory applied to design in hydrology. *Water Resources Research* 8:33–41

De Angelis D, Presanis AM, Conti S, Ades AE. 2014. Estimation of HIV burden through Bayesian evidence synthesis. *Statistical Science* 29:9–17

de Sá TH, Tainio M, Goodman A, Edwards P, Haines A, et al. 2017. Health impact modelling of different travel patterns on physical activity, air pollution and road injuries for São Paulo, Brazil. *Environment International* 108:22–31

Drovandi CC, Holmes C, McGree JM, Mengersen K, Richardson S, Ryan EG. 2017. Principles of experimental design for big data analysis. *Statistical Science* 32:385

Eidsvik J, Mukerji T, Bhattacharjya D. 2015. Value of information in the earth sciences: Integrating spatial modeling and decision analysis. Cambridge University Press

Fang W, Wang Z, Giles MB, Welton NJ, Andrieu C, et al. 2021. Multilevel and quasi Monte Carlo

methods for the calculation of the expected value of partial perfect information. *Medical Decision Making (accepted for publication)*

Felli JC, Hazen GB. 1998. Sensitivity analysis and the expected value of perfect information. *Medical Decision Making* 18:95–109

Friedman JH. 1991. Multivariate adaptive regression splines. *The Annals of Statistics* :1–67

Giles MB, Goda T. 2019. Decision-making under uncertainty: using MLMC for efficient estimation of EVPPI. *Statistics and Computing* 29:739–751

Heath A, Kunst N, Jackson C, Strong M, Alarid-Escudero F, et al. 2020. Calculating the expected value of sample information in practice: Considerations from 3 case studies. *Medical Decision Making* 40:314–326

Heath A, Manolopoulou I, Baio G. 2016. Estimating the expected value of partial perfect information in health economic evaluations using integrated nested Laplace approximation. *Statistics in Medicine* 35:4264–4280

Heath A, Manolopoulou I, Baio G. 2017. A review of methods for analysis of the expected value of information. *Medical Decision Making* 37:747–758

Heath A, Manolopoulou I, Baio G. 2019. Estimating the expected value of sample information across different sample sizes using moment matching and nonlinear regression. *Medical Decision Making* 39:347–359

Hironaka T, Giles MB, Goda T, Thom H. 2020. Multilevel Monte Carlo estimation of the expected value of sample information. *SIAM/ASA Journal on Uncertainty Quantification* 8:1236–1259

Hughes D, Waddingham E, Mt-Isa S, Goginsky A, Chan E, et al. 2016. Recommendations for benefit-risk assessment methodologies and visual representations. *Pharmacoepidemiology and Drug Safety* 25:251–262

Ishizaka A, Nemery P. 2013. Multi-criteria decision analysis: methods and software. John Wiley & Sons

Jackson CH, Bojke L, Thompson SG, Claxton K, Sharples LD. 2011. A framework for addressing structural uncertainty in decision models. *Medical Decision Making* 31:662–674

Jackson CH, Presanis AM, Conti S, De Angelis D. 2019. Value of Information: sensitivity analysis and research prioritisation in Bayesian evidence synthesis. *Journal of the American Statistical Association* 114:1436–1449

Jalal H, Alarid-Escudero F. 2018. A gaussian approximation approach for value of information analysis. *Medical Decision Making* 38:174–188

Johnson R, Woodcock J, de Nazelle A, de Sa T, Goel R, et al. 2019. A guide to value of information methods for prioritising research in health impact modelling. *arXiv preprint arXiv:1905.00008*

Jones M, Goldstein M, Jonathan P, Randell F. 2016. Bayes linear analysis for bayesian optimal experimental design. *Journal of Statistical Planning and Inference* 171:115–129

Koffijberg H, Rothery C, Chalkidou K, Grutters J. 2018. Value of information choices that influence estimates: a systematic review of prevailing considerations. *Medical Decision Making* 38:888–900

Kunst N, Wilson E, Glynn D, Alarid-Escudero F, Baio G, et al. 2020. Computing the expected value of sample information efficiently: practical guidance and recommendations for four model-based methods. *Value in Health* 23:734–742

Laber EB, Meyer NJ, Reich BJ, Pacifici K, Collazo JA, Drake JM. 2018. Optimal treatment allocations in space and time for on-line control of an emerging infectious disease. *Journal of the Royal Statistical Society. Series C, Applied statistics* 67:743

Lindley DV. 1956. On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics* :986–1005

Madan J, Ades AE, Price M, Maitland K, Jemutai J, et al. 2014. Strategies for efficient computation of the expected value of partial perfect information. *Medical Decision Making* 34:327–342

Manski CF. 2019. Treatment choice with trial data: Statistical decision theory should supplant hypothesis testing. *The American Statistician* 73:296–304

Menzies NA. 2016. An efficient estimator for the expected value of sample information. *Medical*

Decision Making 36:308–320
Milborrow S. 2011. earth: Multivariate adaptive regression splines. R package. Derived from mda:mars by T. Hastie and R. Tibshirani.
Morio J. 2011. Global and local sensitivity analysis methods for a physical system. *European Journal of Physics* 32:1577
Mueller N, Rojas-Rueda D, Cole-Hunter T, De Nazelle A, Dons E, et al. 2015. Health impact assessment of active transportation: a systematic review. *Preventive Medicine* 76:103–114
Mytton OT, Jackson C, Steinacher A, Goodman A, Langenberg C, et al. 2018. The current and potential health benefits of the National Health Service Health Check cardiovascular disease prevention programme in England: a microsimulation study. *PLoS Medicine* 15:e1002517
National Institute for Health and Care Excellence. 2013. Guide to the methods of technology appraisal 2013. London: National Institute for Health and Care Excellence
Nutt DJ, King LA, Phillips LD. 2010. Drug harms in the UK: a multicriteria decision analysis. *The Lancet* 376:1558–1565
Oakley JE, Brennan A, Tappenden P, Chilcott J. 2010. Simulation sample sizes for Monte Carlo partial EVPI calculations. *Journal of Health Economics* 29:468–477
Oakley JE, O'Hagan A. 2004. Probabilistic sensitivity analysis of complex models: a Bayesian approach. *Journal of the Royal Statistical Society, Series B* 66:751–769
O'Hagan A, Buck CE, Daneshkhah A, Eiser JR, Garthwaite PH, et al. 2006. Uncertain Judgements: Eliciting Experts' Probabilities. John Wiley & Sons
O'Hagan A, Oakley JE. 2004. Probability is perfect, but we can't elicit it perfectly. *Reliability Engineering & System Safety* 85:239–248
Parmigiani G, Inoue L. 2009. Decision theory: principles and approaches. Wiley
Philips Z, Claxton K, Palmer S. 2008. The half-life of truth: what are appropriate time horizons for research decisions? *Medical Decision Making* 28:287–299
Quinonero-Candela J, Rasmussen CE. 2005. A unifying view of sparse approximate Gaussian process regression. *Journal of Machine Learning Research* 6:1939–1959
Raiffa H, Schlaifer R. 1961. Applied statistical decision theory. Harvard University
Razavi S, Jakeman A, Saltelli A, Prieur C, Iooss B, et al. 2021. The future of sensitivity analysis: An essential discipline for systems modeling and policy support. *Environmental Modelling & Software* 137:104954
Reeves BC, Rooshenas L, Macefield RC, Woodward M, Welton NJ, et al. 2019. Three wound-dressing strategies to reduce surgical site infection after abdominal surgery: the Bluebelle feasibility study and pilot RCT. *Health Technology Assessment* 23:1
Repo AJ. 1989. The value of information: Approaches in economics, accounting, and management science. *Journal of the American Society for Information Science* 40:68–85
Roos M, Martins TG, Held L, Rue H. 2015. Sensitivity analysis for Bayesian hierarchical models. *Bayesian Analysis* 10:321–349
Ryan EG, Drovandi CC, McGree JM, Pettitt AN. 2016. A review of modern computational algorithms for Bayesian optimal design. *International Statistical Review* 84:128–154
Sadatsafavi M, Bansback N, Zafari Z, Najafzadeh M, Marra C. 2013. Need for speed: an efficient algorithm for calculation of single-parameter expected value of partial perfect information. *Value in Health* 16:438–448
Saltelli A, Chan K, Scott EM, eds. 2000. Sensitivity analysis. Wiley
Sobol IM. 1993. Sensitivity analysis for non-linear mathematical models. *Mathematical Modelling and Computational Experiment* 1:407–414
Spiegelhalter DJ, Abrams KR, Myles JP. 2004. Bayesian approaches to clinical trials and health-care evaluation, vol. 13. John Wiley & Sons
Stanfill B, Mielenz H, Clifford D, Thorburn P. 2015. Simple approach to emulating complex computer models for global sensitivity analysis. *Environmental Modelling & Software* 74:140–155
Strong M, Oakley JE. 2013. An efficient method for computing single-parameter partial expected

value of perfect information. *Medical Decision Making* 33:755–766

Strong M, Oakley JE. 2014. When is a model good enough? Deriving the expected value of model improvement via specifying internal model discrepancies. *SIAM/ASA Journal on Uncertainty Quantification* 2:106–125

Strong M, Oakley JE, Brennan A. 2014. Estimating multiparameter partial expected value of perfect information from a probabilistic sensitivity analysis sample: A nonparametric regression approach. *Medical Decision Making* 34:311–326

Strong M, Oakley JE, Brennan A, Breeze P. 2015. Estimating the expected value of sample information using the probabilistic sensitivity analysis sample: A fast, nonparametric regression-based method. *Medical Decision Making* 35:570–83

Thom HHZ, Jackson CH, Welton NJ, Sharples LD. 2017. Using parameter constraints to choose state structures in cost-effectiveness modelling. *Pharmacoeconomics* 35:951–962

Villar SS, Bowden J, Wason J. 2015. Multi-armed bandit models for the optimal design of clinical trials: benefits and challenges. *Statistical Science* 30:199

Walker S, Sculpher M, Claxton K, Palmer S. 2012. Coverage with evidence development, only in research, risk sharing, or patient access scheme? A framework for coverage decisions. *Value in Health* 15:570–579

Welton N, Ades A, Caldwell D, Peters T. 2008. Research prioritization based on expected value of partial perfect information: a case-study on interventions to increase uptake of breast cancer screening. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 171:807–841

Welton N, Strong M, Jackson C, Baio G. 2020. Bayesian Methods in Pharmaceutical Research, chap. Economic Evaluation and Cost-Effectiveness of Health Care Interventions. CRC Press, 301–315

Welton NJ, Sutton AJ, Cooper N, Abrams KR, Ades AE. 2012. Evidence Synthesis for Decision Making in Healthcare, vol. 132. John Wiley & Sons

Willan AR, Pinto EM. 2005. The value of information and optimal clinical trial design. *Statistics in Medicine* 24:1791–1806

Wilson ECF. 2015. A practical guide to value of information analysis. *Pharmacoeconomics* 33:105–121

Wood SN. 2017. Generalized Additive Models: An Introduction with R. CRC Press, 2nd ed.

Woods DC, Overstall AM, Adamou M, Waite TW. 2017. Bayesian design of experiments for generalized linear models and dimensional analysis with industrial and scientific application. *Quality Engineering* 29:91–103