Interpreting the expected value of perfect information about parameters

Susan Griffin and Karl Claxton, Centre for Health Economics, University of York

Abstract

Aims: To investigate the expected value of partial perfect information (EVPPI) and the research decisions it can address.

Methods: Expected value of information (EVI) analysis assesses the expected gain in net benefit from further research. Where the expected value of perfect information (EVPI) exceeds the costs of additional research EVPPI can be used to identify parameters that contribute most to the EVPI and parameters with no EVPPI that may be disregarded as targets for further research. Recently it was noted that parameters with zero EVPPI for a one-off research design may be associated with high EVPPI when considered as part of a sequential design. This paper examines the characteristics and role of conditional and sequential EVPPI in EVI analysis.

Results: We illustrate the calculation of EVPPI for single parameters, groups of parameters, conditional and sequential EVPPI. Conditional EVPPI is the value of perfect information about one parameter, conditional on having obtained perfect information about another. Sequential EVPPI is the value of perfect information for a sequential research design to investigate first one parameter, then another. Conditional EVPPI differs from the EVPPI for a single parameter in the current evidence position. Sequential EVPPI includes elements from the joint EVPPI for the parameters and the EVPPI for the first parameter in sequence. Sequential designs allow abandonment of research on the second parameter on the basis of additional information obtained on the first.

Conclusions: The research decision space addressed by EVI analyses can be widened by incorporating sequential EVPPI to assess sequential research designs.

Introduction

In cost-effectiveness analysis the expected net benefits (NB) of a set of mutually exclusive health technologies are compared in order to identify the intervention that would represent the best use of available resources. The net benefits of each alternative will be estimated with uncertainty. It has been argued that inference is irrelevant to the decision of which technology to adopt on the basis of current evidence and that information about decision uncertainty should inform questions about further research.(Claxton 1999) To this end expected value of information (EVI) analysis can be used to estimate whether there is an expected gain in net benefit to be made by obtaining further information to inform the adoption decision. This paper will examine in more detail the way in which EVI analysis can be used to evaluate research decisions.

The expected value of perfect information (EVPI) describes the value of further research that would eliminate all of the parameter uncertainty in the decision problem. The expected value of partial perfect information (EVPPI) describes the value of obtaining additional information on a subset of input parameters, that is eliminating uncertainty in only some aspects of the decision problem. It has been suggested that EVPPI can be used to identify those parameters that contribute most to the EVPI for the whole decision problem and that might be worthy of further research (ref). Additionally it has been suggested that those parameters or groups of parameters associated with low or no EVPPI may be disregarded as potential targets for further research (ref). It has been emphasised that the sum of the EVPPI for each parameter in the model need not sum to the EVPI for the decision problem as a whole, and that combinations of parameters may be associated with positive EVPPI were they to be investigated jointly even if individually they all exhibit zero EVPPI.(Briggs et al. 2006)

Existing EVI analyses (refs) have focussed on question about one-off research designs that can be used to address the following research decisions:

i) Do not conduct further research
ii) Conduct research to inform a single parameter in isolation?
iii) Conduct research to inform a group of parameters simultaneously?
iv) Conduct research to inform all the parameters simultaneousely, i.e. the whole decision problem.

We take each in turn to illustrate how EVI analysis can inform these research decisions. We will then expand the research decision space to include sequential designs and show how EVI analysis can be used to inform a more compete set of research decisions commonly faced.

One-off research designs

i) is there value in any further research?

With current information a decision maker should select the intervention (j) that maximises expected net benefit. The net benefit of each intervention can be calculated as a function of a set of input parameters which are known with uncertainty. Given a decision problem with two uncertain parameters, $\theta 1$ and $\theta 2$, the value of a decision based on current information can be expressed as:

(1) $\max_j E_{\theta 1,\theta 2} NB(j,\theta 1,\theta 2)$

With perfect information the decision-maker could select the intervention that maximises the net benefit for particular values of θ1and θ2 ($\max_j NB(j,\theta1,\theta2)$). However, the true values of θ1and θ2 are unknown, so the expected value of a decision taken with perfect information is found by averaging the maximum net benefit over the joint distribution of θ1and θ2:

(2) $E_{\theta1,\theta2} \max_j NB(j,\theta1,\theta2)$

The value of a decision made with perfect information (2) can then be compared to the value of the same decision based on current information (1) in order to determine the expected value of perfect information (EVPI):

(3) $EVPI = E_{\theta1,\theta2} \max_j NB(j,\theta1,\theta2) - \max_j E_{\theta1,\theta2} NB(j,\theta1,\theta2)$

Additional information only has value when it would lead to the adoption of a different intervention to the one we would have selected on the basis of current information. The information gained from further research is non-rival and can be used to inform every future decision. Thus the EVPI should be multiplied by the future population of patients expected to benefit from the reduction in decision uncertainty (pop). The EVPI represents the upper bound for the value that could be gained from further research. Where the population EVPI is positive it must be compared to the potential costs of further research (C) in order to establish whether there may be a positive payoff (Π). The costs of research include not only investigative and organisational costs, but also any opportunity cost to patients included in the research. For example, if the additional research were to include randomisation between alternative treatments, then a proportion of patients would not receive the intervention with the maximum expected net benefit on the basis of current evidence. Also those included in research will not be able to benefit from the results, i.e the population that can benefit is used up in the research itself. For there to be any value to further research the maximum achievable value must at least exceed the minimum expected cost of research.

(4) $\Pi = EVPI*pop - \min(C_{\theta1},C_{\theta2})$

Where this first necessary condition is met we may then go on to calculate the value of obtaining further information for a range of more specific research questions. If the expected payoff is negative then investing in further research would not represent a good use of available resources, as any value gained from the additional information would be outweighed by the cost of obtaining that information.

ii) Is there value in further research to inform a single parameter in isolation?
A study could be set up to gather information about one specific input parameter. With perfect information about θ1 the decision maker could select treatment j that maximises the expected net benefit over the remaining uncertain parameter θ2. As the true values of θ1 are unknown this must then be averaged over the distribution of θ1:

(5) $E_{\theta1} \max_j E_{\theta2|\theta1} NB(j,\theta1,\theta2)$

The value of a decision made with perfect information about θ2 (and no further information about θ1) is:

(6) $E_{\theta 2} \max_j E_{\theta 1|\theta 2} NB(j,\theta 1,\theta 2)$

By comparing the value of a decision made with perfect information about either θ1 (5) or θ2 (6) to the value of a decision made with current information (1) we can quantify the expected value of further research to inform either parameter in isolation.

(7) $EVPPI_{\theta 1} = E_{\theta 1} \max_j E_{\theta 2|\theta 1} NB(j,\theta 1,\theta 2) - \max_j E_{\theta 1,\theta 2} NB(j,\theta 1,\theta 2)$

(8) $EVPPI_{\theta 2} = E_{\theta 2} \max_j E_{\theta 1|\theta 2} NB(j,\theta 1,\theta 2) - \max_j E_{\theta 1,\theta 2} NB(j,\theta 1,\theta 2)$

The EVPPI about each parameter must be multiplied by the population of patients who stand to benefit from the additional information, and compared to the potential costs of research to determine whether there may be a positive payoff from that research.

(9) $\Pi_{\theta 1} = EVPPI_{\theta 1} * pop - C_{\theta 1}$

(10) $\Pi_{\theta 2} = EVPPI_{\theta 2} * pop - C_{\theta 2}$

iii) Is there value in further research to inform groups of parameters simultaneously?
A future study could include more than one input parameter so that the results could reduce the uncertainty in a number of inputs simultaneously. In the case of a two parameter model, obtaining perfect information on both θ1 and θ2 would eliminate all the decision uncertainty, and so the joint EVPPI for θ1 and θ2 is equivalent to the EVPI ($EVPPI_{\theta 1,\theta 2} = EVPI$). In order to establish whether there may be a positive payoff from research to update both parameters simultaneously the joint EVPPI should be multiplied by the population expected to benefit and compared to the costs of research:

(11) $\Pi_{\theta 1,\theta 2} = EVPPI_{\theta 1,\theta 2} * pop - C_{\theta 1,\theta 2}$

This also answers question (iv) "is there value in further research to inform the whole decision problem?". But more generally, that question can be informed by evaluating

(12) $\Pi_{\theta 1,...,\theta n} = EVPI * pop - C_{\theta 1,..,\theta n}$

for the total number (N=1,..,n) of uncertain parameters.

Summary
EVI analyses are based on an assessment of the value of a decision based on current evidence (1) and the value of that same decision based on additional evidence about all of the input parameters (2), individual input parameters ((5) and (6)) or groups of parameters (in this example equivalent to (1)). These elements of EVI analysis can be combined to answer a variety of decisions regarding one-off research designs, that is whether there would be value in conducting any single additional study. In order to inform these research decisions it is necessary to estimate the size of the patient population expected to benefit from any additional information and the costs of

conducting the additional research, including any opportunity costs to the participants of that research. If the payoff from an additional study is expected to be positive then investing in further research may represent a valuable use of current resources.

Sequential research designs

Decisions about one-off research designs cover only a portion of the total research decision space. We may wish to expand the research decision space to consider the value of a series of studies, each investigating a different parameter or group of parameters in turn. This gives us a new question not previously addressed in EVI analyses:

(v) Conduct research to inform a set of parameters sequentially?

Recently it has been highlighted that additional information about one input parameter will modify the EVPPI associated with remaining input parameters.(Welton et al. 2006) This would indicate that the value of a sequential research design to investigate first $\theta 1$ then $\theta 2$ cannot be evaluated simply on the basis of $EVPPI_{\theta 1}$ (5) and $EVPPI_{\theta 2}$ (6). What is required is an estimate of the value of obtaining additional information on $\theta 2$, conditional on having already obtained additional information on $\theta 1$.

Conditional EVPPI

If research is conducted to obtain additional information on $\theta 2$ this, in combination with the information already collected for $\theta 1$ provides the decision-maker with perfect information on both parameters. Thus the value of a decision taken with perfect information on $\theta 2$ and $\theta 1$ is simply (2), regardless of whether this information was obtained sequentially or simultaneously.

Again we must compare the value of the decision taken with additional information to the value of the same decision made with current information. Under this premise, current information includes the information already collected on $\theta 1$. Therefore the decision based on current information is based on perfect knowledge of $\theta 1$, and its value is given by (5). Thus the conditional EVPPI for $\theta 2$ given perfect information about $\theta 1$ can be expressed as:

(13) $EVPPI_{\theta 2|\theta 1} = E_{\theta 1,\theta 2} \max_j NB(j,\theta 1,\theta 2) - E_{\theta 1} \max_j E_{\theta 2|\theta 1} NB(j,\theta 1,\theta 2)$

Or we could calculate the conditional EVPPI for $\theta 1$ given perfect information about $\theta 2$:

(14) $EVPPI_{\theta 1|\theta 2} = E_{\theta 1,\theta 2} \max_j NB(j,\theta 1,\theta 2) - E_{\theta 2} \max_j E_{\theta 1|\theta 2} NB(j,\theta 1,\theta 2)$

These values can be multiplied by the population expected to benefit from the additional information and compared to the cost of research to determine the payoff from additional research on a second parameter conditional on having already conducted further research on the first.

(14) $\Pi_{\theta 2|\theta 1} = EVPPI_{\theta 2|\theta 1} * pop - C_{\theta 2}$

(16) $\Pi_{\theta 1|\theta 2} = EVPPI_{\theta 1|\theta 2} * pop - C_{\theta 1}$

This, however, would not answer the question regarding the value of a sequential research design because it does not incorporate the value or cost of the information obtained on the first parameter in sequence.

Sequential EVPPI

The total value of a sequential design can be obtained by combining the value of the information gained on the first parameter in sequence with the conditional EVPPI for the second parameter in sequence. For a sequence evaluating first $\theta 1$ then $\theta 2$ we would sum $EVPPI_{\theta 1}$ (5) and $EVPPI_{\theta 2|\theta 1}$(13). But:

(17) $EVPPI_{\theta 1} + EVPPI_{\theta 2|\theta 1} = E_{\theta 1} \max_j E_{\theta 2|\theta 1} NB(j,\theta 1,\theta 2) - \max_j E_{\theta 1,\theta 2} NB(j,\theta 1,\theta 2) + E_{\theta 1,\theta 2} \max_j NB(j,\theta 1,\theta 2) - E_{\theta 1} \max_j E_{\theta 2|\theta 1} NB(j,\theta 1,\theta 2)$

$= E_{\theta 1,\theta 2} \max_j NB(j,\theta 1,\theta 2) - \max_j E_{\theta 1,\theta 2} NB(j,\theta 1,\theta 2)$

$= EVPI$

The sum of a series of EVPPI calculations conditional on the previously investigated parameters in sequence is equal to the joint EVPPI for that set of parameters. So is the value of a sequential design simply equivalent to the EVPPI for the same group of parameters? The answer is no only if you allow for a change in research decision conditional on the new information as it is collected. So for particular resolutions of $\theta 1$ there may be no expected value from obtaining additional information on $\theta 2$ ($EVPPI_{\theta 2|\theta 1}=0$) and further research should not be conducted. In these circumstances the sequential research design allows the decision maker to forgo additional research on $\theta 2$, i.e. to issue a 'stop' decision. For particular resolutions of $\theta 1$ where $EVPPI_{\theta 2|\theta 1}>0$ the decision may be to 'go' and conduct research on to $\theta 2$. The choice set for the decision maker now includes the intervention (j) and the decision about whether to proceed with research on to the next parameter in sequence (r).

(18) $seqEVPPI_{\theta 1,\theta 2} = E_{\theta 1,\theta 2} \max_r \{\max_j NB(j,\theta 1,\theta 2), \max_j E_{\theta 2|\theta 1} NB(j,\theta 1,\theta 2)\} - \max_j E_{\theta 1,\theta 2} NB(j,\theta 1,\theta 2)$

If research was costless we would obtain any information that reduced the probability of making an incorrect decision. Without considering the cost of research equation (18) would again be equivalent to the EVPI (3). In a sequential design the cost of research into the first parameter in the sequence is always incurred, for example $\theta 1$. However, research on the second parameter in the sequence, for example $\theta 2$, will only proceed and costs be incurred if the expected payoff is positive, for example if $\Pi_{\theta 2|\theta 1} > 0$.

(19) $\Pi_{seq\theta 1,\theta 2} = E_{\theta 1,\theta 2} \max_r \{(\max_j NB(j,\theta 1,\theta 2) - \max_j E_{\theta 2|\theta 1} NB(j,\theta 1,\theta 2))*pop - C_{\theta 2}, 0\} + (E_{\theta 1} \max_j E_{\theta 2|\theta 1} NB(j,\theta 1,\theta 2) - \max_j E_{\theta 1,\theta 2} NB(j,\theta 1,\theta 2))*pop - C_{\theta 1}$

$= E_{\theta 1,\theta 2} \max_r \{(\max_j NB(j,\theta 1,\theta 2) - \max_j E_{\theta 2|\theta 1} NB(j,\theta 1,\theta 2))*pop - C_{\theta 2}, 0\} + \Pi_{\theta 1}$

(20) $\Pi_{seq\theta 2,\theta 1} = E_{\theta 1,\theta 2} \max_r \{(\max_j NB(j,\theta 1,\theta 2) - \max_j E_{\theta 1|\theta 1} NB(j,\theta 1,\theta 2))*pop - C_{\theta 1}, 0\} + \Pi_{\theta 2}$

Equations (19) and (20) provide answers question (v): "is there value in further research to inform a set of parameters sequentially?". They describe the expected

maxium payoff from a sequential design before any of the parameters have been investigated further. They are composed of the same elements (equations (1), (2), (5) and (6)) that are used to answer questions (i) to (iv), alongside some assessment of the costs of research. Therefore expanding the research decision space to incorporate questions about sequential research designs should not increase the complexity of future EVI analyses.

Sequential EVPPI can be used to identify a parameter or group of parameters that should be initiated first as it evaluates the potential impact of any additional information on any parameters to be investigated subsequently. Once the information has been collected on the first parameter in sequence then the previous EVI calculations become redundant and a new set incorporating the additional information is required. In this example, once additional information has been collected for $\theta 1$, the question of whether to proceed with research on $\theta 2$ will be informed by an new calculation of $\Pi_{\theta 2}$ (10).

Comparisons between alternative research designs

For the two-parameter decision problem used to illustrate the EVI calculations it would be feasible to enumerate the payoff from all of the possible one-off and sequential research designs. The decision-maker could then select the approach that would maximise the payoff from further research, including consideration of the option for no further research. However, as the number of parameters increases, the number of potential research designs also increases.

Comparing benefit

The value of a sequential design will be at least as great as the value of information about the first parameter in the sequence. Furthermore, the value of the sequential design cannot exceed the value of information for both parameters in combination. So, in general:

(21) $\quad EVPPI_{\theta 1} \leq seqEVPPI_{\theta 1,\theta 2} \leq EVPI_{\theta 1,\theta 2}$

Where the payoff from investigating the second parameter in sequence is always negative, the sequential design collapses back to the EVPPI for the first parameter in sequence. Where the payoff from investigating the second parameter in sequence is always positive the value of information gathered from a sequential design will be equal to the value obtained from a one-off research design investigating both parameters simultaneously.

The conditional EVPPI for a parameter can be identified solely from an assessment of one-off research designs because:

(22) $\quad EVPPI_{\theta 2|\theta 1} = EVPI_{\theta 1,\theta 2} - EVPPI_{\theta 1}$

Where the EVPPI and the conditional EVPPI for a parameter are equal to zero (i.e. $EVPPI_{\theta 1} = EVPPI_{\theta 1|\theta 2} = 0$) then we may disregarded it as a potential target for future research.

Comparing costs

The total cost of investigating two parameters sequentially will in general exceed the cost of investigating them simultaneously. This may in part be due to economies realised by running a single trial to investigate both parameters. However, even in the absence of any economies from a joint design, the opportunity costs of the sequential design will exceed the opportunity costs of the joint design because the length of time taken to gain the additional information on both parameters will be greater for a sequential design.

(22) $C_{\theta 1} + C_{\theta 2} > C_{\theta 1,\theta 2}$

This would indicate that where the payoff from research into the second parameter in sequence is always positive a sequential design will be dominated by a one-off research design that investigates both parameters simultaneously. The value in the sequential approach derives from the ability to avoid the cost of research on the second parameter in sequence on the basis of additional information collected on the fist parameter in sequence. This suggests that the benefits may be most valuable when the costs of investigating the second parameter in the sequence is significant and the payoff for investigating the second parameter in sequence varies widely with the value of the first parameter in sequence and there is a possibility that research about the second will be unnecessary.

Discussion

Previous EVI analyses have focussed on questions about one-off research designs. When assessing the individual EVPPI for each parameter in the model, it may be tempting to prioritise further studies in order of the size of the individual payoffs. In this paper we have discussed how to calculate the payoff from alternative sequential research designs that implicitly accounts for learning from the information gathered on each parameter in sequence in terms of determining whether to proceed with research on the next parameter in sequence. Widening the research decision space to incorporate sequential designs is important because the sequence with the largest expected payoff may imply a different order of research to that indicated by naively ordering the parameters in terms of the individual EVPPIs.

This paper focuses on perfect information to describe the alternative research decisions. Of course the value of perfect information represents the maximum value of further research, and in practice we will obtain only imperfect or sample information. Even so, estimating the maximum payoff from further research is computationally feasible and can provide a useful indication of which research designs are likely to represent a worthwhile use of available resources. Currently, optimising over all possible research design space on the basis of the expected value of sample information is generally computationally expensive and maybe prohibitive in many common circumstances. Therefore using this type of EVPI analysis to prioritise the research decision space is of value and can be used to focus expensive EVSI calculation on those research designs which are likely to be most valuable.

References (to be completed)

Briggs A, Sculpher M and Claxton K Decision modelling for health economic evaluation. Oxford, Oxford University Press. 2006.

Claxton K The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. J Health Econ 1999 **18**: 342-64.

Welton NJ, Ades AE, Caldwell DM, et al. Research Prioritisation Based on Expected Value of Partial Perfect Information: a Case Study on Interventions to Increase Uptake of Breast Cancer Screening. Submitted to JRSS A: Statistics in Society as a Discussion paper 2006.