
Article:
Tuffaha, H.W., Strong, M. orcid.org/0000-0003-1486-8233, Gordon, L.G. et al. (1 more author) (2016) Efficient Value of Information Calculation Using a Nonparametric Regression Approach: An Applied Perspective. Value in Health, 19 (4). pp. 505-509. ISSN 1098-3015

https://doi.org/10.1016/j.jval.2016.01.011

Efficient value of information calculation using a non-parametric regression approach: an applied perspective

Haitham W Tuffaha[*1,2], Mark Strong[3], Louisa G Gordon[1,2], Paul A Scuffham[1,2]

Affiliations

1. Menzies Health Institute Queensland, Griffith University, Gold Coast, QLD Australia.
2. Centre for Applied Health Economics, School of Medicine, Griffith University, Meadowbrook QLD, Australia.
3. School of Health and related Research (ScHARR), University of Sheffield, Sheffield, UK.

***Corresponding Author contact details**

Centre for Applied Health Economics, School of Medicine, Griffith University, QLD, Australia

haitham.tuffaha@griffith.edu.au

Tel: 61 7 338 21510

Fax: 61 7 338 21338

Key words

Value of information, non-parametric regression, Monte Carlo simulation

Acknowledgment

Haitham Tuffaha is supported by a National Health and Medical Research Council PhD scholarship through the Centre for Research Excellence in Nursing Interventions for

Hospitalised Patients.

Mark Strong is supported by a post-doctoral National Institute for Health Research fellowship. This report is independent research supported by the UK National Institute for Health Research (Post-Doctoral Fellowship, PDF-2012-05-258). The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health.

Conflict of interest

The authors declare no conflict of interest.

Efficient value of information calculation using a non-parametric regression approach: an applied perspective

Abstract

Value of information (VOI) analysis provides an analytical framework to assess whether obtaining additional evidence is worthwhile to reduce decision uncertainty. The reporting of VOI measures, particularly, the expected value of perfect parameter information (EVPPI) and the expected value of sample information (EVSI), is limited because of the computational burden associated with typical two-level Monte Carlo-based solution. Recently, a non-parametric regression approach was proposed that allows the estimation of multi-parameter EVPPI and EVSI directly from a probabilistic sensitivity analysis sample. We used the regression approach to calculate EVPPI and EVSI in two models, and compared the results with the estimates obtained via standard Monte Carlo simulation. VOI values from the two approaches were very close; however, the regression method was faster. We conclude that the non-parametric regression-based approach provides an efficient and easy-to-implement alternative for EVPPI and EVSI calculation in economic models.

1. Introduction

Decision models are commonly used to evaluate the cost-effectiveness of health interventions. They are populated with input parameters estimated from various sources; however, the true values of these parameters are not known with certainty, which may result in suboptimal decisions.(1) The preferred approach to characterise decision uncertainty is to conduct probabilistic sensitivity analysis (PSA) whereby uncertainty is propagated in the model using Monte Carlo simulation.(2) Decision uncertainty is then presented as the probability that each intervention has the highest expected net benefit (i.e., benefits minus costs). Nevertheless, an important additional step is to know whether a decision can be made based on current evidence or if additional research is required. This can be informed using value of information (VOI) analysis.(3) Measures of VOI include (1) the expected value of perfect information (EVPI), which is the maximum value of additional information to resolve all uncertainty in the parameters; (2) the expected value of perfect parameter information (EVPPI), which is the value of resolving uncertainty in a given parameter or set of parameters; and (3) the expected value of sample information (EVSI), which estimates the value of a particular data collection exercise (e.g. a randomised controlled trial with some chosen sample size) in reducing decision uncertainty.(4)

EVPI calculation is straightforward given the PSA; however, although this measure is necessary, it is not sufficient to inform decisions because it represents only an upper-bound of the value of additional research to resolve uncertainty.(3) Rather, it is important to know which parameters are contributing most to decision uncertainty, such that further research should focus on these. Here, the EVPPI for some parameter represents the value of eliminating uncertainty about that parameter, and therefore gives an upper-bound on the value of a study to inform that parameter. The EVSI meanwhile represents the value of a

given study design in reducing parameter uncertainty.(5) Comparing the EVSI with the expected cost of a research study establishes a sufficient condition to inform whether additional research is worthwhile. Unfortunately, the reporting of EVSI and EVPPI estimates in economic evaluations remains limited because of the perceived computational burden associated with these two measures.(6-8)

The EVPPI for a single parameter or a group of parameters is typically calculated using a two-level nested Monte Carlo simulation approach. This requires sampling values of the parameter(s) of interest in an outer loop, and then conditional on each sampled parameter set, sampling from the joint conditional distribution of the remaining parameters in an inner loop. At each inner loop step the model is evaluated.(9, 10) For EVSI calculation, plausible sets of data from the proposed future study of a given sample size are simulated in an outer loop, and then conditional on each generated data set, the posterior distribution of the parameters is sampled in an inner loop. Again, the model is evaluated at each inner loop step.(11, 12) The repeated sampling and evaluation of the model within the inner loop is time consuming. Calculating EVSI values for a range of possible sample sizes could take days or even weeks depending on the complexity of the model.(12-14) Furthermore, it is often difficult to determine the number of simulations at each level to ensure adequate precision and to avoid the upward bias that results from the maximisation over sampled quantities that occurs within the outer loop of the simulation.(15) Finally, the presence of parameter correlation or non-conjugacy between prior parameter distributions and proposed data likelihoods makes EVSI calculation even more difficult.(11) Here, Markov Chain Monte Carlo simulation (or some similar approach) will be necessary.(6, 7) In some situations, most notably in multi-linear (i.e. sum-product type) models (e.g., decision tree) where the net benefit is a linear function of the cost and effect parameters, or when the incremental net

benefit is approximately normally distributed, one-level Monte Carlo simulation or analytical equations can be used.(11, 16, 17) However, there is a wide class of models for which these constraints do not apply.

Methods for efficient EVPPI calculation of single parameters have been developed. These show promise, but do not extend to groups of parameters simultaneously.(18, 19) A method based on the numerical approximation of the posterior expected net benefit, conditional on sampled data, has been proposed as an efficient approach for EVSI calculation; however, it requires significant skills and effort to write the necessary computer code.(13, 20) Recently, Strong et al. have proposed a more straightforward non-parametric regression approach for calculating multi-parameter EVPPI and EVSI directly from a PSA sample.(9, 21) The method has the advantage in that the model does not need to be run as part of the EVPPI or EVSI algorithm. Nevertheless, there is a need to demonstrate the value of this method in real-world cases and to compare its performance with the standard approach of Monte Carlo simulation.

In this paper we apply the non-parametric regression method to calculate the EVPPI and the EVSI in two decision models for two healthcare interventions. In addition, we compare the results and computation time with the estimates obtained using Monte Carlo simulation.

2. Methods

2.1 The two economic models

We conducted two cost-effectiveness analyses using two decision models constructed in TreeAge Pro (Version 2014 R1). The full details of the two models and analyses can be found elsewhere.(22, 23)

Model 1: Negative pressure wound therapy in caesarean section patients

The first model was a decision tree for negative pressure wound therapy (NPWT) compared with hydrocolloid dressing in preventing surgical site infections following caesarean sections in high-risk (e.g., obese) women.(23) The modelled patients may develop surgical site infection which could be either superficial or deep. Patients could die or survive depending on the type of the infection developed (Appendix 1). To populate the model we systematically searched the literature and identified relevant evidence. Due to the scarcity of information on the effectiveness of NPWT in this setting, we combined the data on the relative effectiveness of the device from a pilot study (n=92) on obese women undergoing caesarean sections with the data from a trial (n=81) on NPWT in high-risk patients with various types of surgeries.(23)

Model 2: Nutritional support for the prevention of pressure ulcers in hospitalized patients

The second model was a six-health-state Markov cohort model for nutritional support compared with standard hospital diet in preventing pressure ulcers.(22) Model duration was one year with a one-day cycle length. Patients start the model with intact skin before they move sequentially between different stages of skin ulceration (i.e., closed wound to open wound). Further, patients could either die of any cause, be discharged, or remain hospitalised (Appendix 1). We systematically searched and identified relevant evidence. We performed a meta-analysis of five trials (n=1,381) to estimate the relative effectiveness of nutritional support in preventing pressure ulcers compared with hospital diet.(22)

The two models were probabilistic; input parameters were assigned probability distributions. In general, beta distributions for probabilities and utilities, gamma distributions for costs and disutilities, and lognormal distributions for relative risks.(22, 23) For the set of

unknown input parameters (θ), each model predicted the net benefit (NB) for each intervention (i), thus NB (i, θ) = willingness-to-pay* Effect (i, θ) – *Cost (i,* θ). The efficacy outcome in the two models was quality-adjusted life-years (QALYs) gained, and we set the willingness-to-pay threshold at $50,000/QALY. The preferred intervention would be the one with the maximum expected NB ($\max_i E_\theta$ NB (i, θ)). In each case, a PSA was performed using Monte Carlo simulation (10,000 iterations) to characterise decision uncertainty.

2.2 Value of information calculation

We calculated VOI measures using the standard Monte Carlo and the Strong et al. non-parametric regression approach for each of the two decision models. We also recorded, for each decision problem, the computation time for each VOI approach.

Methods to calculate value of information measures using Monte Carlo simulations are described in detail elsewhere.(11, 24, 25) In summary, we started our analysis by calculating the EVPI, which is the difference between the expected NB of a decision with perfect information and the decision based on current information:(3)

$$EVPI = E_\theta \max_i NB(i, \theta) - \max_i E_\theta NB(i, \theta) \qquad \text{Equation 1}$$

The EVPPI for the parameter(s) of interest θ_I is the difference between the expected NB with perfect information on these parameters, conditional on the complementary set of other parameters θ_C, and the expected NB with current information: (5, 25)

$$\boldsymbol{EVPPI_{\theta_I} = E_{\theta_I} max_i E_{(\theta_C|\theta_I)} NB(i, \theta_I, \theta_C) - max_i E_\theta NB(i, \theta)} \qquad \text{Equation 2}$$

For the pressure ulcer Markov model, we performed two nested Monte Carlo simulation procedures with 1,000 simulations in each loop. We found this number of simulations sufficient for the estimates to converge.(25) We assumed the NPWT model is linear with no

correlation between input parameters, and therefore, a one-level simulation scheme was used in which we sampled from θ_I, but kept θ_C fixed at their prior mean.(25)

The EVSI is the difference between the expected value of a decision made after collecting data (D) on the parameter(s) of interest and the expected NB with current information:(11)

$$\boldsymbol{EVSI = E_D max_i E_{\theta_C,\theta_I|\,D}\, NB(i,\theta_I,\theta_C) \; - \; max_i E_\theta NB(i,\theta)} \qquad \text{Equation 3}$$

We assumed the prior distribution and data were conjugate distributions, and thus, the posterior distribution was known in closed form.(11, 14) We conducted a two-level Monte Carlo simulation with 1,000 iterations in each of the inner and outer loops for the Markov pressure ulcer model.(11) For the NPWT decision tree, we avoided nested simulation in estimating the NBs by "plugging in" the prior means of θ_C and the posterior means of θ_I.(11)

We repeated EVPPI and EVSI calculations for the two models using regression methods in R software as described by Strong et al. (9, 21) Using the PSA sample of 10,000 iterations (K =10,000), we fitted a regression model for each decision option. After that we extracted the regression model fitted values denoted as $\hat{g}\left(i,\theta_I^k\right)$ and calculated EVPPI via: (9)

$$\boldsymbol{EVPPI_{\theta_I} \;\cong\; \frac{1}{K}\sum_{k=1}^{K} \; max_i\, \hat{g}\left(i,\theta^k{}_I\right) - \; max_i \; \frac{1}{K}\sum_{k=1}^{K} \; \hat{g}\left(i,\theta_I^k\right)} \text{ Equation 4}$$

To calculate EVSI, we generated data and calculated summary statistic D^k conditional on each sample θ_I^k in the PSA. For the relative risk parameter for example, we calculated the summary statistics by generating a sample data of the probability of the event in the intervention group (P_i^k) from a Binomial (P_i^k, n) and for the control group (P_c^k) from a Binomial (P_c^k, n); thus, D^k = log (P_i^k/P_c^k). Then we fitted a regression model for each

decision option and extracted the regression model fitted values denoted as $\hat{g}(i, D^k)$. EVSI was estimated via:(21)

$$\boldsymbol{EVSI} \cong \frac{\mathbf{1}}{\boldsymbol{K}}\sum_{\boldsymbol{k=1}}^{\boldsymbol{K}} \boldsymbol{max}_{\boldsymbol{i}}\, \hat{\boldsymbol{g}}\left(\boldsymbol{i}, \boldsymbol{D}^{\boldsymbol{k}}{}_{\boldsymbol{I}}\right) - \boldsymbol{max}_{\boldsymbol{i}}\, \frac{\mathbf{1}}{\boldsymbol{K}}\sum_{\boldsymbol{k=1}}^{\boldsymbol{K}} \hat{\boldsymbol{g}}\left(\boldsymbol{i}, \boldsymbol{D}^{\boldsymbol{k}}{}_{\mathbf{I}}\right) \quad \text{Equation 5}$$

3. Results

The incremental NB of NPWT was $70 with 65% probability being cost-effective. The per-person EVPI was $76. The parameter with the highest EVPPI was the relative risk of surgical site infection. The per-person EVPPI was estimated at $75 using Monte Carlo simulation, and $74 using non-parametric regression. The per-person EVSI for a future study to inform this parameter was $63 when calculated with Monte Carlo simulation compared with $61 using regression. The calculation time for the EVPPI and EVSI measure was short (around one minute) in both Monte Carlo simulation and non-parametric regression.

For nutritional support intervention, the incremental NB was $675 indicating it is cost-effective. The probability this intervention is cost-effective was 85%, and the per-person EVPI was $33. The parameter with the highest per-person EVPPI value was the relative risk of pressure ulcer which had a value of $17 under both methods of computation. The per-person EVSI of a study to inform this parameter was approximately $6 under both approaches. In terms of the computation time, non-parametric regression estimated EVPPI and EVSI values in less than one minute and in one step. For the Monte Carlo simulation, every EVPPI estimate took around four hours, and every EVSI value for a given sample size took around eight hours. Table 1 summarises the results and Figure 1 illustrates the EVSI curves from the two models.

4. Discussion

This study reports the first comparison of the non-parametric regression and Monte

Carlo methods for EVPPI and EVSI computation in real-life decision problems. The estimates from the two methods were close; however, VOI computation using regression methods was faster, particularly in the Markov model. The regression approach eliminated the need for two-level simulation in the Markov model. Further, we avoided the Bayesian updating step for EVSI calculation. This benefit of the regression method becomes very important in the computation of EVSI when the parameter prior is not conjugate to the data likelihood.

The non-parametric regression method used in this paper represents a significant contribution in resolving the computational burden often perceived with EVPPI and EVSI calculation. The method provides a flexible approach to calculate VOI measures for models built in any software and of any complexity. The approach does not require knowledge of the statistics language R, or sophisticated programming ability. Although the R code is provided, the Sheffield Accelerated Value of Information (SAVI) online tool provides a free platform for VOI calculation.(26) It only requires that the PSA files be uploaded before VOI measures can be calculated and presented. We anticipate that the uptake of this efficient approach will increase as time passes. This will hopefully encourage more reporting of VOI estimates in economic evaluations and encourage a wider adoption of VOI analysis as a useful tool to inform funding decisions and to optimise research design and prioritisation.

It is worth mentioning that another efficient approach has been recently proposed by Jalal and Kuntz for EVSI calculation from a PSA sample.(27) In their method, they use linear regression metamodeling with the assumption that the incremental NB is normally distributed.(27) Unfortunately, the normality assumption, and the assumption of linearity of the model could make it difficult to generalise this approach.(28) However, it would be interesting to see how the two regression approaches, and other efficient methods, compare

using models of various types. It would be useful also to compare the performance of these efficient methods with Monte Carlo simulation using more complex models such as those using microsimulation, or models with non-conjugate priors (e.g., Weibull distribution), or with correlated parameters.

In conclusion, the non-parametric regression-based approach provides an efficient, flexible and easy-to-implement alternative for EVPPI and EVSI calculation in economic models. The approach should facilitate the wider incorporation of VOI analysis in decision frameworks.

References

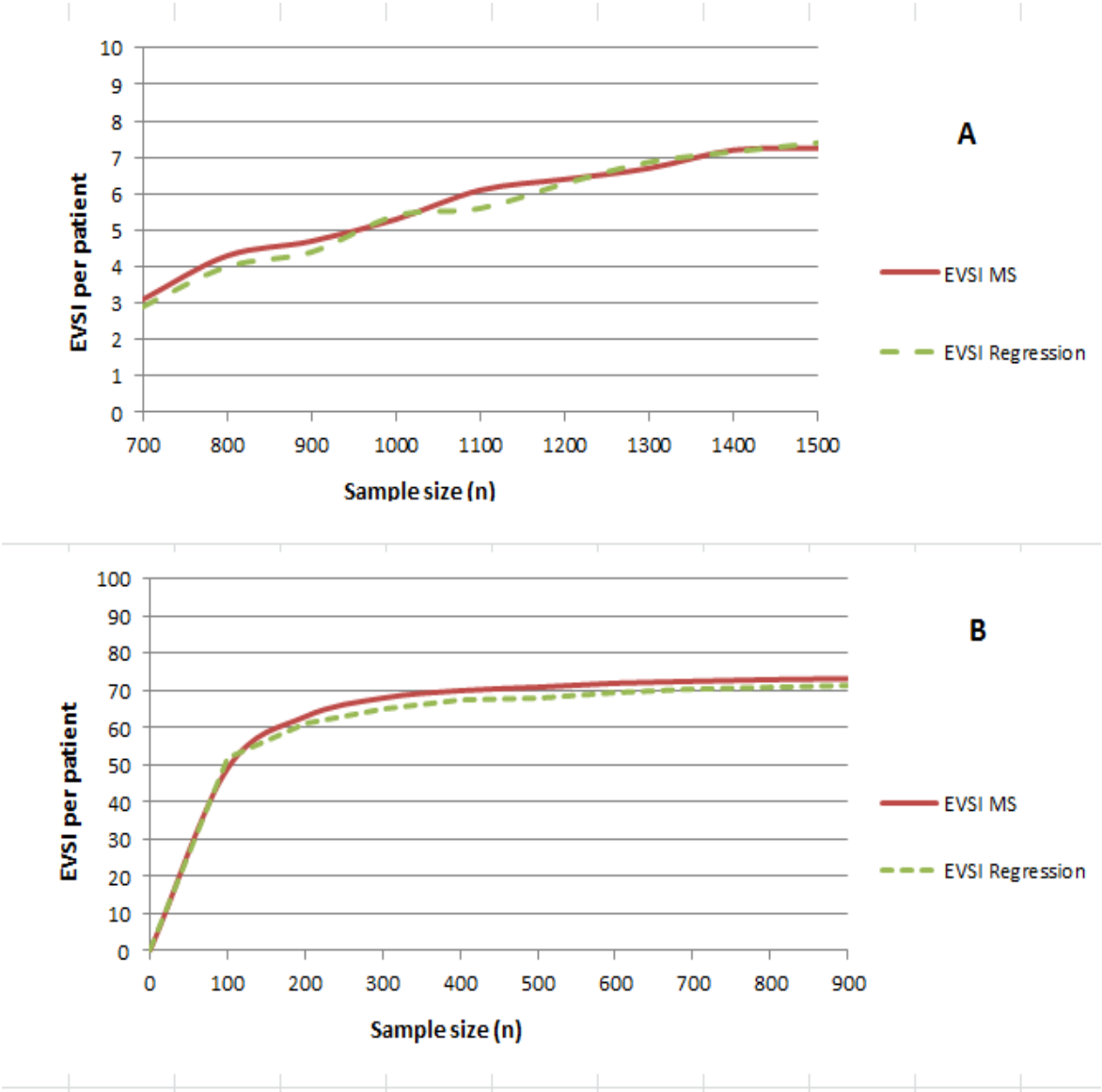
1. Claxton K. Exploring uncertainty in cost-effectiveness analysis. Pharmacoeconomics. 2008; 26: 781-98.
2. Briggs AH, Weinstein MC, Fenwick EAL, et al. Model parameter estimation and uncertainty: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--6. Value Health. 2012; 15: 835-42.
3. Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. J Health Econ. 1999; 18: 341-64.
4. Claxton K, Posnett J. An economic approach to clinical trial design and research priority-setting. Health Econ. 1996; 5: 513-24.
5. Claxton KP, Sculpher MJ. Using value of information analysis to prioritise health research: some lessons from recent UK experience. Pharmacoeconomics. 2006; 24: 1055-68.
6. Steuten L, van de Wetering G, Groothuis-Oudshoorn K, et al. A systematic and critical review of the evolving methods and applications of value of information in academia and practice. Pharmacoeconomics. 2013; 31: 25-48.
7. Tuffaha HW, Gordon LG, Scuffham PA. Value of information analysis in oncology: the value of evidence and evidence of value. Journal of oncology practice / American Society of Clinical Oncology. 2014; 10: e55-62.
8. Tuffaha HW, Gordon LG, Scuffham PA. Value of information analysis in healthcare: a review of principles and applications. Journal of medical economics. 2014; 17: 377-83.
9. Strong M, Oakley JE, Brennan A. Estimating multiparameter partial expected value of perfect information from a probabilistic sensitivity analysis sample: a nonparametric regression approach. Med Decis Making. 2014; 34: 311-26.
10. Brennan A, Kharroubi S, O'Hagan A, et al. Calculating partial expected value of perfect information via Monte Carlo sampling algorithms. Med Decis Making. 2007; 27: 448-70.
11. Ades AE, Lu G, Claxton K. Expected value of sample information calculations in medical decision modeling. Med Decis Making. 2004; 24: 207-27.
12. Strong M, Oakley JE, Brennan A, et al. Estimating the Expected Value of Sample Information Using the Probabilistic Sensitivity Analysis Sample: A Fast Nonparametric

Regression-Based Method. Med Decis Making. 2015.
13. Brennan A, Kharroubi SA. Expected value of sample information for Weibull survival data. Health Econ. 2007; 16: 1205-25.
14. Soares MO, Dumville JC, Ashby RL, et al. Methods to assess cost-effectiveness and value of further research when data are sparse: negative-pressure wound therapy for severe pressure ulcers. Med Decis Making. 2013; 33: 415-36.
15. Oakley JE, Brennan A, Tappenden P, et al. Simulation sample sizes for Monte Carlo partial EVPI calculations. J Health Econ. 2010; 29: 468-77.
16. Willan AR, Pinto EM. The value of information and optimal clinical trial design. Statistics in Medicine. 2005; 24: 1791-806.
17. Coyle D, Oakley J. Estimating the expected value of partial perfect information: a review of methods. The European journal of health economics : HEPAC : health economics in prevention and care. 2008; 9: 251-9.
18. Sadatsafavi M, Bansback N, Zafari Z, et al. Need for Speed: An Efficient Algorithm for Calculation of Single-Parameter Expected Value of Partial Perfect Information. Value Health. 2013.
19. Strong M, Oakley JE. An Efficient Method for Computing Single-Parameter Partial Expected Value of Perfect Information. Med Decis Making. 2012.
20. Brennan A, Kharroubi SA. Efficient computation of partial expected value of sample information using Bayesian approximation. J Health Econ. 2007; 26: 122-48.
21. Strong M, Oakley J, Brennan A, et al. Estimating the Expected Value of Sample Information using the Probabilistic Sensitivity Analysis Sample. A Fast Non-Parametric Regression Based Method. Med Decis Making. 2015.
22. Tuffaha HW, Roberts S, Chaboyer W, et al. Cost-Effectiveness and Value of Information Analysis of Nutritional Support for Preventing Pressure Ulcers in High-risk Patients: Implement Now, Research Later. Appl Health Econ Health Policy. 2015; 13: 167-79.
23. Tuffaha HW, Gillespie BM, Chaboyer W, et al. Cost-utility analysis of negative pressure wound therapy in high-risk cesarean section wounds. The Journal of surgical research. 2015.
24. Claxton K. Bayesian approaches to the value of information: implications for the regulation of new pharmaceuticals. Health Econ. 1999; 8: 269-74.
25. Groot Koerkamp B, Myriam Hunink MG, Stijnen T, et al. Identifying key parameters in cost-effectiveness analysis using value of information: a comparison of methods. Health Econ. 2006; 15: 383-92.
26. Strong M, Breeze P, Thomas C, et al. SAVI - Sheffield Accelerated Value of Information.https://github.com/Sheffield-Accelerated-VoI/SAVI. Release version 2.0.0 (2015-03-09) ed.
27. Jalal H, Goldhaber-Fiebert JD, Kuntz KM. Computing Expected Value of Partial Sample Information from Probabilistic Sensitivity Analysis Using Linear Regression Metamodeling. Med Decis Making. 2015.
28. Welton NJ, Thom HH. Value of Information: We've Got Speed, What More Do We Need? Med Decis Making. 2015.

Table 1: A comparison of value of information measures calculation using Monte Carlo simulation and non-prametric regression

Intervention		EVPPI in $ (SE)	Approximate computation time	EVSI in $ (SE)[c]	Approximate computation time
Monte Carlo simulation[a]					
NPWT	RR site infection:	74.8 (1.6)	1 minute	63.0 (1.3)	1 minute
	Other parameters:	3.0 (0.3)	1 minute		
Nutritional	RR pressure ulcer:	17.4 (2.2)	4 hours	6.4 (0.7)	8 hours
support	Other parameters:	8.9 (1.7)	4 hours		
Non-parametric regression[b]					
NPWT	RR site infection:	74.3 (0.6)	1 minute	61.2 (0.8)	1 minute
	Other parameters:	5.8 (1.4)	1 minute		
Nutritional	RR pressure ulcer:	17.2 (1.0)	1 minute	6.3 (0.8)	1 minute
support	Other parameters:	9.4 (1.0)	1 minute		

[a] Two-level Monte Carlo of 1,000 iterations each in the nutritional support model, and single level simulation of 10,000 iterations in the NPWT model.
[b] From a probabilistic sensitivity sample of 10,000 iterations.
[c] For a sample size of 200 patients in the NPWT model and 1,200 patients in the nutritional support model.
$ = Australian dollar; NPWT = negative pressure wound therapy; EVPI = expected value of perfect information; EVPPI = expected value of perfect parameter information; EVSI = expected value of sample information; RR = relative risk; SE = standard error.



EVSI = Expected value of sample information; MS = Monte Carlo simulation

Figure 1: EVSI curves for the nutritional support model (A) and NPWT model (B)