Assessing Parameter Importance in Decision Models
Application to Health Economic Evaluations

Sandra Milev

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies (FGPS) in partial fulfillment of the requirements for the MSc. degree in Systems Science

Systems Science Program
University of Ottawa, Ottawa, Ontario, Canada

ABSTRACT

Background: Uncertainty in parameters is present in many decision making problems and leads to uncertainty in model predictions. Therefore an analysis of the degree of uncertainty around the model inputs is often needed. Importance analysis involves use of quantitative methods aiming at identifying the contribution of uncertain input model parameters to output uncertainty. Expected value of partial perfect information (EVPPI) measure is a current gold- standard technique for measuring parameters importance in health economics models. The current standard approach of estimating EVPPI through performing double Monte Carlo simulation (MCS) can be associated with a long run time. Objective: To investigate different importance analysis techniques with an aim to find alternative technique with shorter run time that will identify parameters with greatest contribution to parameter uncertainty. *Methods*: A health economics model was updated and served as a tool to implement various importance analysis techniques. Twelve alternative techniques were applied: rank correlation analysis, contribution to variance analysis, mutual information analysis, dominance analysis, regression analysis, analysis of elasticity, ANCOVA, maximum separation distances analysis, sequential bifurcation, double MCS EVPPI,EVPPI-quadrature and EVPPI- single method. *Results*: Among all these techniques, the dominance measure resulted with the closest correlated calibrated scores when compared with EVPPI calibrated scores. Performing a dominance analysis as a screening method to identify subgroup of parameters as candidates for being most important parameters and subsequently only performing EVPPI analysis on the selected parameters will reduce the overall run time.

ACRONYMS AND ABBREVIATIONS

AEC	Actual elasticity coefficient
AROS	Absolute relative overall sensitivity
CEA	Cost effectiveness analysis
CUA	Cost utility analysis
CV	Contribution to variance
DA	Dominance analysis
DM	Dominance measure
DVT	Deep venous thrombosis
EVPPI	Expected value of partial perfect information
EVPI	Expected value of perfect information
EVSI	Expected value of sample information
HTA	Health technology assessment
ICER	Incremental cost effectiveness ratio
ICUR	Incremental cost utility ratio
INB	Incremental net benefit
LY	Life year
NB	Net benefit
MCS	Monte Carlo simulation
MI	Mutual information
MI THR	Minimal invasive total hip replacement
MSD	Maximum separation distance
PE	Pulmonary embolism
PSA	Probabilistic sensitivity analysis
SA	Sensitivity analysis
SB	Sequential bifurcation
SMR	Standardized mortality ratio
SRC	Standardized regression coefficient
STD THR	Standard total hip replacement
THR	Total hip replacement
VIF	Variance inflation factor
VOI	Value of information
QALY	Quality adjusted life year

ACKNOWLEDGEMENTS

My thesis builds on previous work by my supervisor Dr. Doug Coyle. His unreserved support, openness, bright ideas and great sessions made this work a very pleasant experience for me. Dr. Coyle also influenced me to get into the world of health economics and gave a new shape to my career which I absolutely enjoy and I am grateful for. It's been a great privilege and pleasure to work with Dr. Coyle and I hope that this is only the beginning of our collaboration.

To my mom, I am grateful for all the last-minute trips that she made to Ottawa to help me get through my studies easier. More important, I am thankful for her never-ending aim to make me a better human.

To Roberto, I am grateful for his support in this journey of mine, or I should better say, journey of our family. His focus and perseverance are true inspiration.

To my dearest Kalina, I own apologies for all the times that mommy couldn't play because she needed to study. Her warm hugs were solutions to my many frustrations; her scribbles on my textbooks will always remain as very sweet memorabilia.

To all my dear friends, thank you for your support

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ALGORTHMS

Chapter 1

Introduction and Background

The following chapter outlines the objectives and structure of the thesis. Health economic evaluations compare two or more health care interventions in both costs and outcomes. The purpose of health economic evaluations is to provide information on value for money of health care interventions that can lead to a more efficient use of health care resources. Since there are many factors that can contribute to the uncertainty around making the right decisions, importance analysis is used to identify those key parameters that have the highest impact on uncertainty and for which there may be value in further research.

1.1 Introduction

1.1.1 Statement of the Problem

Uncertainty in parameters is present in many risk assessment and decision making problems and leads to uncertainty in model predictions. Therefore an analysis of the degree of uncertainty of the model inputs is often needed. Sensitivity analysis includes various analytic methods that explore the relationship between the changes in the model inputs and outputs. On the other hand, importance analysis involves the use of quantitative methods aiming to identify the contribution of uncertain input model parameters to output uncertainty (Borganovo 2006). Such methods are called uncertainty importance measures, and include ranking of the model input variables by their impact to the model uncertainty. This information helps both in identifying the most important contributors to output uncertainty, but also in finding the most effective ways of reducing that uncertainty, by providing information on which parameters it may be worthwhile obtaining more information (Smith and Croke 2006).

Importance analysis has applications in models where any uncertainty has to be eliminated, such as models assessing nuclear safety. In these cases importance analysis will contribute to reducing the uncertainty in the model by identifying those parameters which can lead to system failure. On the other hand, importance analysis has applications in models where some degree of uncertainty can be tolerated and the purpose of importance analysis is to identify which are parameters that are important in terms of the degree of uncertainty they cause. Thus, the role of importance analysis is to identify parameters for which it might be worthwhile obtaining more information, such as those used in health economic models.

This research will focus on importance analysis whose aim is to identify the contribution of uncertain input parameters to the model output and its specific application to health economics models. Expected value of partial perfect information (EVPPI) is an information-based measure of importance that combines the probability of an incorrect decision making being made with the consequential loss function and is a current gold standard in health economic evaluations (Briggs et al 2012). However, evaluating EVPPI does come with a computational cost (Claxton 2008). Health economic models are often non-linear and complex and calculation of EVPPI may require many simulation runs of the same probabilistic model, which can become a large burden given the time constraints the decision makers often face. Another factor that needs to be considered is that health economic models are predominately built in Microsoft Excel to facilitate easy adoption by a wide range of potential users (Bujkiewicz et al. 2011), and MS Excel is not the most efficient platform for conducting complex analyses relating to the calculation of EVPPI.

1.1.2 Research Objectives

The objectives of this research are:

1. To update an existing health economic model for minimally invasive arthroplasty in management of hip arthritic disease (Coyle et al. 2008) that will serve as a case study.
2. To review current available measures of parameter importance techniques.
3. To apply the identified techniques of importance to the case study.
4. To compare results based on various measures of importance and contrast these with estimates based on the expected value of partial perfect information (EVPPI) measure which is a current gold- standard importance measure in health economic models.
5. Based on objective 4, to determine which importance measure(s) can be used as a screening technique to identify subgroup of parameters with potentially high EVPPI. The

ultimate aim is reserving estimation of EVVPI for those parameters only, which will result with overall decrease in run-time

1.1.3 Relevance to Research

Decision makers need to decide whether to reimburse or not given health technology and one of the decision options should be whether or not to fund additional research. There is always uncertainty surrounding the reimbursement decision and therefore linking the decision to reimburse with the decision to find additional research is often needed. EVPPI is a technique that identifies the maximum worth of further research to reduce uncertainty; however it has been argued that it comes at certain computational expense. By reducing the computational burden of estimating EVPPI with selecting relevant subgroup of parameters, the importance analysis in health economic evaluations might become more accessible and feasible. This may have the additional benefit of encouraging wider adoption of this method in health economic evaluations.

1.1.4 Outline

This chapter provides an introduction to health economic evaluations as well as an introduction to importance analysis. The second chapter provides a description of the health economic model comparing the cost- utility of a minimally invasive arthroplasty in the management of hip arthritic disease (minimal invasive hip replacement) to the standard arthroplasty (standard hip replacement).

The third chapter outlines the techniques for assessing parameter importance and provides an algorithm for the implementation of each of the techniques. Chapter four presents the results of each of the assessed techniques, as well as a comparison with the EVPPI results. The fifth chapter discusses results and provides conclusions on which of the importance measure(s) can be considered appropriate screening technique(s).

1.2 Background

1.2.1 Introduction to Health Economic Evaluations

Health economic evaluations compare two or more health care interventions in both costs and outcomes, with the aim of providing information that can lead to more efficient use of health care resources. Therefore, the purpose of any economic evaluation is to identify, measure and compare the costs and outcomes of the health care alternatives.

With the current increasing health care expenditure that puts at risk the sustainability of health care systems, well-informed decisions need to be made regarding funding of health care interventions. Free markets do not exist in healthcare, and therefore active decisions need to be made about which health services should be funded given the scarce resources available (Fox-Rushby and Cairns 2005).

There are different types of health economic evaluations that can be used. They differ primarily by the outcomes measured, with the most known being cost-effectiveness and cost-utility analysis.

Cost –effectiveness analysis (CEA) is a type of evaluation that measures health in terms of a single outcome measure. The outcome measure in CEA is often so called "natural outcomes", such as life years gained or cases avoided, but it can also be disease specific outcome such as percentage of number of seizures decreased, mm Hg of blood pressures reduced etc. CEA always includes a single effect, clinical or health effect of interest, for both alternatives. Although it can be applicable to many different diseases and different types of health programs, the main limitation of CEA is its focus on a single outcome measure which precludes comparison across disease areas. Also, the results of CEA are often difficult to interpret, since a benchmark for what constitutes cost-effective treatment based on the maximum willingness to pay for a disease specific outcome is often unknown.

Cost-utility analysis (CUA) is a type of economic intervention that includes QALY as a measure of health consequences. QALY is an acronym for Quality Adjusted Life Years, a measure that takes into accounts both the quantity and quality of life generated by healthcare interventions. QALY in essence is an arithmetic product of life expectancy and a measure of the quality of the remaining life-years, and therefore it incorporates any changes in survival duration as well as in health related quality of life. QALYs have cardinal properties, i.e. the differences are quantifiable on an interval scale, can be compared across clinical areas and are meant to reflect individual preferences.

Health economic evaluations are used to help set priorities over which treatments should be funded. If a new intervention results with improved outcomes and increased costs, then the incremental cost per additional unit of outcome should be evaluated. By using the incremental cost per additional unit of outcome, resources can be allocated to maximize the net health benefit for the target population under fixed budget (Detsky and Naglie 1990).

Incremental cost effectiveness ratio (ICER) is a common measure in health economic evaluations. ICER is the ratio of the incremental change in costs to incremental change in benefits of the treatments under evaluation. The equation for ICER is:

$$ICER = \frac{C_1 - C_2}{E_1 - E_2}$$

where C_1 and E_1 are the cost and health outcome in the treatment intervention 1 and where C_2 and E_2 are the cost and health outcome in the treatment intervention 2. Costs are usually described in monetary units while health outcomes can be measured in terms of a single clinical outcome (in CEA) or QALYs gained (in CUA). When outcomes are expressed as QALYs, analysts often refer to ICERs as ICURs; i.e. an incremental cost utility ratio.

The explicit route with CEA and CUA is through the comparison of an incremental cost-effectiveness ratio with a threshold (λ) (Sculpher and Claxton 2012). If a treatment is more effective ($E_1 > E_2$) and the associated $ICER < \lambda$, the intervention is deemed to be cost-effective.

The threshold λ has two distinct interpretations. Firstly, it can be interpreted as the shadow price of a unit of health benefit given a constrained budget (Coyle 2004). For instance, Sculpher and Claxton argue that the threshold should represent the health outcomes forgone due to the displacement of existing services to fund any additional cost of new programmes and technologies (i.e. it should reflect opportunity cost) (Sculpher and Claxton 2012). Secondly, λ can be interpreted as the value society places on a unit of health benefit (Coyle, 2004). Within this research study, the first interpretation of λ is adopted, although that would not have an impact of the methods for analysis of parameter importance.

$50,000 per QALY is a common and widely accepted benchmark for λ used by decision making committees across Canada (Winquist et al 2012), and as such will be used in this research.

ICERs are the most popular method of presenting results of CEA and CUA, and yet they have certain limitations. Drummond argues that the ratio does not give idea of the size or scale of treatments being considered; testing for statistical differences between ratios gives rise to some additional complications; as well as reduction in cost and reduction in outcome can result with same ICER when there is increase in cost and increase in outcome (Drummond et al 2005).

Therefore, an additional summary measure for cost effectiveness is incremental net benefit (INB) which is a function of both incremental costs and effects and the threshold λ. Net benefit of a treatment can be expressed in terms of net health benefits or net monetary benefits. The net monetary benefit for a treatment t (NB_t) defined in monetary terms is simply the difference between the measured outcomes and costs weighted by the value of a unit of health benefit (λ).

$$NB_t t = E_t * \lambda - C_t$$

Conversely, the net health benefit for an individual treatment t is defined as:

$$NHB_t = E_t - C_t/\lambda$$

Since the objective of health economic models is comparison of two or more treatments, we are interested in incremental net benefits (INB).

Incremental net benefit (INB) can be calculated as the difference in net benefits between two treatment choices t_1 and t_2

$$INB_{t_1,t_2} = \lambda * \left(E_{t_1} - E_{t_2}\right) - \left(C_{t_1} - C_{t_2}\right) = NB_{t_1} - NB_{t_2}$$

INB is equal to zero when the incremental cost-effectiveness ratio of the treatment being studied versus the comparator treatment is equal to the threshold. By definition, INB will be positive for the comparison between the optimal treatment and any other potential treatment option (Coyle 2004).

Although, the estimation of an ICER is seen to be a standard outcome for economic evaluation by the majority of health economists working in this field, some have questioned the adoption of a single value for λ (Birch and Gafni, 2006). Decision makers could adopt values for λ that vary by the context of the disease or intervention under consideration. However, this would not impact the appropriateness of the methods discussed for measuring parameter importance – only the estimate of net benefit to be considered.

Estimation of the cost effectiveness of healthcare interventions is typically achieved with use of decision-analytical modeling. Analytical models as a primary tool in health economics have a range of uses including the synthesis of data from various sources and extrapolation from primary data sources (Briggs and Sculpher 1998).

The range of modeling techniques for medical and economic decision modeling has advanced substantially lately, although the relative simplicity of cohort-based models is still an attraction for many modelers and decision makers (Caro 2012). Cohort-based models or aggregate models are type of models that examine the proportions of the population undergoing different events with associated costs and outcomes. Two most common types of cohort-based models used in health economic evaluation are decision trees and Markov models.

A decision tree is a decision supporting tool that uses a tree-like graph of decisions and their possible consequences, including chance of event outcomes and resource costs. It flows from left to right beginning with an initial clinical choice on a defined category of patients (Drummond et al 2005). Decision trees outline the probability and the valuation of each outcome, such as a QALY, cost or net benefit measure. The expected value of each outcome for each alternative treatment option is computed analytically ("rollback") by summing the probability of each outcome with its value (Brennan et al 2006).

Complex decision trees could be adopted which include many decision strategies, differential timing of events and many multiple outcomes from chance nodes. However, there are limitations with decision trees in many contexts. First, the structure of a decision tree allows patients to progress through the model in one way only, which can be problematic for conditions with recurrent states. Secondly, it is difficult within a decision tree to incorporate a temporal element without making the tree unduly complex. Therefore, those elements of economic modeling that are time dependent can be difficult to implement (Drummond et al 2005). Finally, decision trees can become very complex when modeling long-term implications, particularly related to chronic diseases.

The second type of analytical cohort models commonly used in health economic assessments is Markov models. Markov models are used in other disciplines to represent stochastic processes that are random processes which evolve over time. In the field of medical decision analysis, they are particularly suited to modeling the progression of chronic disease (Briggs and Sculpher 1998). In essence, Markov models are based on series of "health states" that a patient can enter at any given time. Time elapses explicitly in Markov models and transitions probabilities are assigned for movement between these states over a discrete time period known as a 'Markov cycles".

By attaching estimates of resource use and health outcome consequence to the states and transitions in the model, and then running the model over a large number of cycles, it is possible to estimate the long term costs and outcomes associated with a disease and a particular healthcare intervention (Briggs and Sculpher 1998).

Markov models can more easily handle the issues associated with decision trees aforementioned and are therefore better for representing more complex processes occurring over time, such as long term complications associated with chronic diseases.

The main limitation of Markov models is its memoryless feature, or also known as "Markovian assumption". In a Markov model, transitional probability of one state to another is independent of the earlier transitions, i.e. the model cannot remember where the patient came from; that is it cannot distinguish the origin of patients in a state at a given time point and they are considered as homogenous group (Drummond et al 2005). However, this limitation can be reduced by creating additional health states related to the patient history, i.e. "tunnel states". The states of the cycles can be accessed only in a pre-determined logical sequence simulating patients' health history. For example, if modeling schizophrenia, it might be relevant whether relapsers are relapsing for the first time, or they have experienced a previous relapse. Therefore, instead of relapse as a health state, we would create first-time relapse, second-time relapse and relapse with more than 2 previous relapses.

Besides the cohort-based models, a different analytical approach exists, where the decision problem takes into account the extensive history of the individuals with their specific characteristics. These patient i.e. individual level models are known as discrete events simulations (DES) or “microsimulations”, which can record the patient’s individual baseline characteristics, how they change over time and take into account historical factors such as prior health states or interventions (Roberts et al 2012). DES, instead of modeling events one cycle at a time, simulate sequences of events by drawing directly probability distributions of event times (Weinstein, 2006).

Infectious disease modeling is a further patient-level modeling approach that can handle interaction between individuals as well as changes in cohort size and disease characteristics over time.

For the purpose of this research, individual level models and their specific characteristics will not be considered.

So far, the discussion has focussed on analysis conducted within a deterministic framework, where the incremental costs and benefits of the interventions under evaluation are known with certainty. However, the input parameters that go into the model are subject to some uncertainty that should be reflected in the cost-effectiveness results as well. Probabilistic methods can be used to describe the uncertainty associated with input parameters and, in turn, can estimate the uncertainty in the model outputs of cost, effect, and cost-effectiveness (Briggs, 2000). This can be achieved by implementing an informal Bayesian approach to cost-effectiveness analysis, specifying relevant parameters as probability distributions rather than point estimates. Specifically, for those input parameters that could be estimated from observed data, consideration should be given to the prior distribution of these parameters to reflect uncertainty. This technique allows the estimation of the likelihood of various output values based on a wide number of sets of input parameters generated by sampling from their probability density functions In specifying the nature of the uncertainty around input parameters, care should be taken to ensure that the prior distributions chosen are consistent with any logical bounds on the parameter values (Briggs, 2000). This approach of fitting distribution to parameters based on the available evidence indeed does not represent the formal Bayesian approach which includes an application of a likelihood function in an effort to combine new observed data with the prior distribution. However, this informal Bayesian approach has been widely accepted in the field of health economics. For mathematical convenience, it is common for prior distributions to be specified in terms of a distribution that is conjugate to the likelihood function based on the observed data since the use of a conjugate prior leads to a posterior distribution from the same family of distributions (Briggs 1999). It has also been argued that this approach would lead to very similar results if a formal Bayesian with uninformative priors is implemented (Briggs et al 2006).

Within the stochastic framework, the cost-effectiveness results would be based on expected values for cost and health benefits (e.g. QALYs). Expectations are based on data obtained from the Monte Carlo simulation (MCS), i.e. running the model more than once (for example 5,000 times) to allow sampling from the various distributions.

If $E(C_1)$ and $E(C_2)$ are the expected value of health benefits (e.g. QALYs) from treatment 1 and treatment 2 respectively, and $E(E_1)$ and $E(E_2)$ are the expected values of health benefits, then based on the output from a MCS, the ICER is the ratio of the expected values of incremental costs and incremental benefits; not the expected value of the ratio (Coyle et al 2003).

$$ICER = \frac{E(C_{1)} - E(C_2)}{E(E_1) - E(E_2)}$$

Coyle argued that decision makers should solely be concerned with the expected value of outcomes when basing decisions over the optimal treatment choice (Coyle et al 2003). Thus, in considering uncertainty it is necessary to obtain measures of expected values and the recommended approach is that of probabilistic analysis based on Monte Carlo simulation (Coyle et al 2003). This does not, however, mean that further consideration of the uncertainty over a decision is irrelevant. Rather, further focus when handling uncertainty should be put on what further information it is justified to collect (Coyle et al 2003). Therefore, conducting importance analysis to identify those parameters that have the highest impact on the final results is a natural next step.

1.2.2 Introduction to Importance Analysis

Decision makers need to decide both whether or not to reimburse a given health technology and the related decision of whether or not to fund additional research. There is always uncertainty surrounding the reimbursement decision and the decision to reimburse is closely linked with the decision to fund additional research. Therefore, it is necessary to simultaneously address both questions “is there sufficient evidence to reimburse” and “is further research for health technology assessment potentially worthwhile”?

A previous US Secretary of Defense noted that “now that we have begun to use economic evaluation in decision making, we have moved from a position of (largely) ‘unknown unknowns’ to one where we have many more ‘knowns’, but also many ‘known unknowns’ (Drummond and Schulper 2006).

However, Bettonvil and Kleijnen pointed out that frequently there are only a small number of key factors which are important, even when the total number of factors is large (Bettonvil and Kleijnen 1996). Therefore, the purpose of importance analysis is to identify those key parameters that have the highest impact and that potentially would be a question of further research.

The model output may respond to parameters in two ways: either the model results may be highly correlated with an input parameter so that small changes in the input value result in significant changes in the model output, or the variability, or uncertainty associated with an input parameter can be propagated through the model resulting in a large contribution to the output variability (Hamby 1994).

Importance analysis involves the use of techniques to determine how different input parameters contribute to the uncertainty over the outcomes of interest (Coyle et al 2003). Importance measures quantify the contribution that individual parameters have to the overall uncertainty. These measures provide information on which parameters it may be worthwhile obtaining more information, although only certain techniques can actually assess the value for such information (Coyle et al 2003).

It is important to make the clear distinction between sensitivity analysis and importance analysis. While the purpose of sensitivity analysis is to address the uncertainty around the model output dependent on the input parameters, importance analysis deals with quantifying input parameters' contribution to the uncertainty.

Briggs made a clear categorization of the different parameters in a health economic model (Briggs 2000). He made distinction between different 'levels' of parameters: (i) parameters relating to analytical methods (e.g. the discount rate, time horizon) employed in an evaluation; (ii) parameters that describe the characteristics of a patient sample (e.g. age/gender) ; and (iii) parameters that could, in principle, be sampled if an appropriate study were designed to collect the relevant data, such as transition probabilities for movement between states over time in a Markov model, resource use consequences and health outcome consequences of the programmes under evaluation; and cost information for resource use (Briggs 2000).

Importance analysis will deal only with the third group of parameter, the reason being that the parameters relating to analytical methods need to be predetermined by following the principles of good modeling approach, while the parameters that describe the characteristics of a patient sample must be determined by the clinical context.

Chapter 2

Case Study (Health Economic Model)

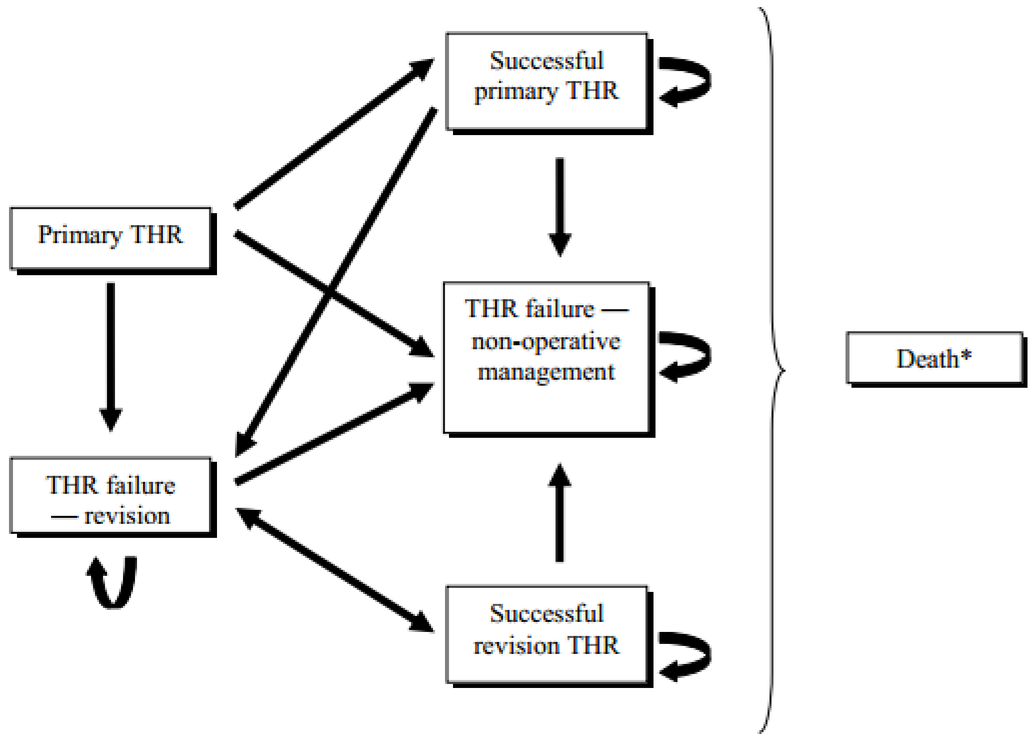
For the purpose of this thesis, different measures of importance were applied to a health economic model comparing the cost- utility of minimally invasive arthroplasty in the management of hip arthritic disease (minimal invasive hip replacement) to the standard arthroplasty (standard hip replacement).

In this chapter the health economic model will be described, its detailed model structure along with the data feeding the model.

2.1. Background

The original model was developed by Coyle et al. (2008) with a purpose of assessing the cost-effectiveness of minimal invasive total hip replacement (MI THR) to a standard total hip replacement (Standard THR) (Coyle et al 2008). This model used Markov simulation to estimate the long term costs and quality adjusted life years for patients undergoing MI and standard THR on a time horizon of 40 years, using one year cycle length. (Figure 1)

Figure 1: Simplified Markov Model used in the Original Model by Coyle et al, 2008

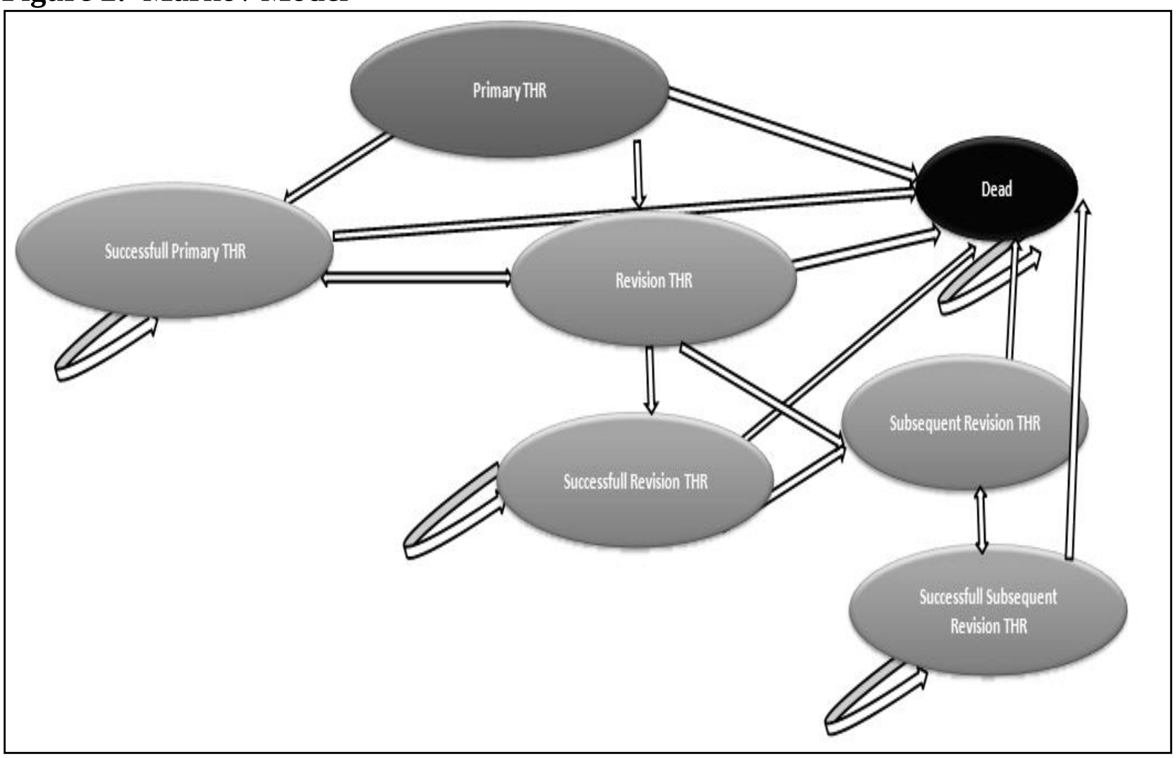


***All states in the model may lead to death**
Source: Coyle D, Coyle K, Vale L, Verteuil R, Imamura M, Glazener C, Zhu S, CADTH Technology Report, HTA, (2008) Minimally Invasive Arthroplasty in the Management of Hip Arthritic Disease: Systematic Review and Economic Evaluation

2.2 Model Structure

Similar to the original model, the newly developed model is a Markov model estimating the long term costs and quality adjusted life years for patients undergoing MI and standard THR. Results are presented in terms of estimates of costs and QALYs per patient. The model uses a time horizon of 40 years, with one year cycle length. All patients that entered the model underwent a total hip replacement, either minimal invasive or standard. The model structure was modified from the original one, such that one year after the surgery, patients are assumed to enter one of the following health states: (1) successful THR, (2) failed THR requiring primary revision, (3) successful primary revision, (4) failed primary revision requiring subsequent revision, (5) successful subsequent revision and (6) death.

Figure 2: Markov Model



In contrast to the original model, the updated version of the model assumes no option of non-operative management due to THR failure. This assumption is based on a new literature review supporting the assumption that following THR failure all patients now undergo a THR revision.

The entry stare in the model is Primary THR. From here, a patient can move to one of three states – successful primary THR, revision THR or death. From successful primary THR a patient can move to revision THR or death or remain in the current state.

From THR revision a patient can require a subsequent revision, die or move to the health state of successful revision THR. Similarly from the successful THR revision state a patient can move to the subsequent THR revision state, die or remain in the health state of successful revision THR. For subsequent revision the movements in health states is as above. All health states can lead to the absorbing state, i.e. death.

The proportion of patients in each health state is therefore weighted with the respective health utilities and costs. The probabilities of the patients moving from one health state, i.e., the transitional probabilities are outlined in section 2.3.1 below.

First, a deterministic analysis was undertaken, where data were analyzed as point estimates and results were presented in the form of ICURs. A probabilistic analysis was subsequently performed in which all uncertain model input parameters were assigned probability distributions. This approach allowed more comprehensive consideration of the uncertainty characterizing the input parameters and captured the non-linearity characterizing the economic model structure. Moreover, the relationship between input parameters and the ICUR is multiplicative, since probabilities and relative risks are used to be determined the proportion of patients in each health state, which are weighted with their utilities and costs.

Results of the probabilistic analysis were summarized in the form of cost effectiveness acceptability curves, which express the probability of each intervention being cost effective at various levels of willingness-to-pay per QALY gained (that is, at various cost effectiveness thresholds).

The model was developed in Microsoft Excel (version 2007, Microsoft Corporation, Redmond, WA, USA), with macros developed in Microsoft Visual Basic for Applications (VBA) (version 2007, Microsoft Corporation, Redmond, WA, USA). Although it might be argued that MS Excel might not be the best choice for stochasting modeling, it is still predominant software tool for conducting health economic evaluations. It is frequently adopted as it facilitates sharing models with decision makers and other researchers without the need to have and understand alternative software options.

2.3 Model Inputs

All data inputs in the model were represented by their mean and standard error. Representing the variability of the data inputs by their standard errors, i.e. standard deviation of the population mean as opposed to the standard deviation of the sample set was felt to be appropriate since interest is in the uncertainty around the estimates of the mean of the population rather than the variability of the data points in the sample set.

2.3.1 Transition Probabilities

The transition probabilities for the model, where available, were taken directly from the literature. All patients are assumed to enter the model upon undergoing minimal invasive or standard THR.

As patients transitioned through the 40 years of the model, Canadian annual rates of mortality were applied to the study population (Life Tables, Statistics Canada, 2002). A weighted average of mortality for men and women was used to derive the rates of mortality used in the model.

All surgical procedures requiring general anesthetic are subject to risk of death during the operation. Therefore, a transition probability of 0.0013 for operational mortality was used in the base case analysis. The operative mortality estimate was based on the results from a retrospective study (Parvizi et al 2001). This study reported 0.013% (4/30,714) intraoperative deaths, each of them associated with events known to produce embolization of marrow contents.

Lie and al. have studied the mortality after THR of almost 40,000 patients with a mean age of 69 years (Lie et al 2004). They observed a lower mortality in patients with THR than in the general population, with a standardized mortality ratio (SMR) of 0.81. SMR represents the ratio of the observed patient mortality and the mortality of the general population with corresponding age, gender and year of birth. As the authors of this study suggest, the reduced mortality in the patients that underwent THR is most likely result of preoperative selection of healthier people. Since the entry age of patients on the model corresponds to the age of the patients in Lie study (Lie et al 2004), this SMR ratio of 0.81 was applied in the model.

In the original model (Coyle et al 2008) the annual rate of revision rate of revision was derived from the Swedish registry, with the risk of revision increasing from one year to the next until year 24. After year 24, it remained constant for the 40-year duration of the model. In the new model that is being used, the annual risk of first revision was fixed and equal to 0.01134, while the annual risk of subsequent revision, if the first one failed, was equal to 0.02751. These probabilities were derived from a more recent source (Lie et al 2004), suggesting that the annual risk of subsequent revisions stays constant over time.

In the original model, four complications were considered: deep venous thrombosis (DVT), pulmonary embolism (PE), dislocation, and deep infection. For all four complications, it is assumed that patients would receive treatment and, upon resolution, would continue to be considered to have undergone successful THRs. With the revised model, infection and dislocation were assumed to be a combined complication based on the available data. The table below summarizes the rates of complications used in the model.

Table 1: Incidence of Complications with THR Included in the Model

Complication	Annual Rate	95% CI		Source
DVT	0.0085	0.0040	0.0130	Eriksson, 2008
PE	0.0026	0.0010	0.0060	Eriksson, 2008
Deep Infection/Dislocation	0.00004	N/A	N/A	Engesaeter, 2003
CI= confidence interval; DVT= Deep venous thrombosis; PE=pulmonary embolism; THR= total hip replacement				

2.3.2 Relative Risks

There are three complications that are considered in the model, dislocation/deep infection, deep venous thrombosis (DVT) and pulmonary embolism (PE). As a result of the paucity of data in the systematic review, the authors of the original model decided that calculated risks were generally implausible because the confidence intervals (CI) surrounding the values were unreasonably large (Coyle et al 2008). Therefore, the original model assumed no difference between MI and Standard procedures and long-term outcomes and implemented constant relative risk of one for revision as well for the four complications.

As a part of this thesis, a search for new data on the relative risks was conducted to explore the possibility of implementing new relative risks that will closely reflect the difference between the procedures. However, the new literature search did not yield any results other than those included in the Coyle study. In order to account for uncertainty around the relative risks, lognormal distribution with a ln mean of 0 and a standard error of 0.5 was incorporated.

2.3.3 Utilities

The utilities used in the model were derived using data from a Canadian clinical trial (Charles et al 2006). Data related to responses to the SF-36 questionnaire from which utility values were derived based on the work by Brazier and al study (Brazier et al 2002). Data in this study was collected from 40 participants. In order to address the issue of missing data for number of individuals, there were three studies conducted, each of them using different methodology approach. The first method used the last value carried forward, excluding the participant who did not have data for the baseline assessment. The second approach data for all participants were included, regardless of its completeness. In the final analyses, those cases that did not have a complete data set were not included, which left only 24 cases within the analyses. As reported, in each of these three analyses mean scores of each time point were compared in a covariance analyses with adjustments made for base line SF6D scores.

These three different approaches did not yield significantly different results. Therefore the results from the first method, i.e. the analysis using last value carried forward for missing values were used in the model.

The utility values for first year THR, successful primary THR, THR revision and successful THR revision for both standard and MI THR were derived using the utility values from table 2, and are represented in table 3 below.

Table 2: Health Utility Values and their Probability Distribution Parameters

Parameter	Utility	Standard Error	Distribution	Distribution Parameters
Standard THR Utilities				
Pre-operative	0.6527	0.0355	Beta	alpha=117, beta=62
3 months	0.7632	0.0242	Beta	alpha=233, beta=72
6 months	0.8014	0.0223	Beta	alpha=255, beta=63
1 year	0.8139	0.0281	Beta	alpha=155, beta=35
Incremental utility for MI THR				
3 months Incremental	0.045	0.0310	Normal	mean=0.045, se=0.031
6 months Incremental	0.001	0.0300	Normal	mean=0.001, se=0.030
1 year Incremental	-0.011	0.0330	Normal	mean=-0.011,se=0.033
Other utility values				
Failure of THR or complications	0.3200	0.0150	Beta	alpha=309, beta=657
Successful Revision	0.6200	0.0150	Beta	alpha=649, beta=398

Table 3: Derived Health States Utility Values

Health State	Formula	Utility Value
First Year – Std THR	0.125**U_preop*+0.25**U_3mth*+0.375**U_6mth*+ 0.25**U_12mth*	0.7764
Success Primary- Std THR	(1-*P_dvt*-*P_pe*-*P_reopinf*)* *U_12mth* +(*P_dvt*+*P_pe*+*P_reopinf*)**U_fail*	0.8089
First Year- MI THR	0.125**U_preop*+0.25*(*U_3mth+UI_3mth)*+0.375*(*U_6mth+UI_6mth)*+0.25*(*U_12mth+UI_12mth)*	0.7853
Success Primary- MI THR	(1-*P_dvt*-*P_pe*-*P_reopinf*)* (*U_12mth+UI_12mth)* +(*P_dvt*+*P_pe*+*P_reopinf*)**U_fail*	0.7980
Revision THR	0.5**U_fail*+0.5**U_rev*	0.4683

IU_3mth=incremental utility at 3month for MI THR; *IU_6mth*=incremental utility at 6 months for MI THR, *IU_12mth*=incremental utility at 12 month for MITHR; MI THR= minimal invasive total hip replacement; *P_dvt*= probability of dvt; *P_pe*=probability of pulmonary embolism; *P_reopinf*= probability of reoperation due to infection; STD THR= standard total hip replacement ; *U_3mth*=utility at 3month for std THR; *U_6mth*=utility at 6 months for stdTHR, *U_12mth*=utility at 12 month for std THR; *U_fail* = utility for THR failure; *U_preop*=preoperative utility; *U_rev*=utility for revision THR

The difference in utilities among STD THR and MI THR at 3 months is clinically significant. As a comparison, acute myocardial infarction is associated with disutility of 0.04 (Sullivan and Ghushchyan 2006) However, the difference in utilities among STD THR and MI THR at 6 months and onwards drastically attenuates which speaks to the short term benefit of MI THR.

2.3.4 Resource Use and Costs

The original model was developed in 2008 and therefore the costs incorporated in the model reflected the actual Canadian cost in 2008. This updated version of the model uses the cost information based on the most recent available cost sources, such as Ontario

Schedule of Benefits and Fees (OHIP Schedule of Benefits and Fees 2010) and Ontario Case Costing Initiative (OCCI 2010).

The costs of the original operation were estimated from the hospital costs associated with each arm of a Canadian clinical trial that compares STD and MI THR (Charles et al 2006). These costs were originally obtained from Ottawa Hospital case-costing system, which includes details on patient specific inpatient costs by categories, allied health, pharmaceuticals, operation room, equipment and non-physician personnel. Mean cost for both MI and STD THR were calculated and inflated with Canadian core CPI to reflect 2010 Canadian costs.

Physician costs were included by incorporating costs for operations and ward visits, based on Ontario Schedule of Benefits and Fees (OHIP Schedule of Benefits and Fees 2010).

As in the original model, all revisions were assumed to be standard THR, because they were considered to be too complex for MI technique. The costs of complications were also captured in the model, using Ontario Case Costing Initiative (OCCI 2010) as a data source.

The tables below summarize the costs for MI, standard THR and cost of complications.

Table 4: Breakdown of Costs Associated with Standard THR

Standard THR Cost		
Procedure	Mean	Std. Error
Allied Health	$537.53	$52.05

Standard THR Cost		
Procedure	Mean	Std. Error
Endoscopy	-	-
Imaging	$ 57.37	$3.19
Lab	$ 150.85	$22.31
Nursing	$2,684.47	$156.16
OR and recovery room	$ 2,637.73	$117.92
Implants	$3,258.12	$144.47
Medications	$ 252.83	$25.50
Ward visits	$245.00	$122.50
Surgical fee	$942.00	$471.00
Total	$10,765.89	$547.61*

*Standard error of the Total is calculated as square root of the sums of squares of the standard error of each of the costs

Table 5: Breakdown of Costs Associated with MI THR

MI THR Cost		
	Mean	Standard Error
Allied Health	$824.36	$134.91
Endoscopy	$18.06	$18.06
Imaging	$125.35	$49.93
Lab	$179.53	$ 58.43
Nursing	$3,240.06	$594.90
OR and recovery room	$2,893.75	$159.35
Implants	$3,384.54	$ 229.46
Medications	$290.01	$44.62
Ward visits	$263.00	-
Surgical fee	$ 942.00	-
Total	$12,160.66	$677.03*

*Standard error of the Total is calculated as square root of the sums of squares of the standard error of each of the costs

Table 6: Costs of Complications and other Costs associated with THR

Description	Mean	Std Error
Cost for re-operation due to dislocation	$7,776	$1,614
Cost of DVT for those patients readmitted as inpatient	$6,600	$424
Cost of PE	$7,442	$555
Cost of re-operation due to infection	$24,377	$1,966
Cost of Std THR operation and first year follow-up cost	$10,766	n/a
Cost of MI THR operation and first year follow up	$12,161	n/a
Cost of follow up cost for successful standard, MI or revision THR	$45	n/a
Cost of revision	$16,199	$1,098

DVT= deep vein thrombosis; PE= pulmonary embolism; Std THR= standard total hip replacement, MI THR= minimal invasive total hip replacement

2.4 Base Results

The deterministic analysis showed that standard THR is less costly than minimal invasive THR, with lifetime costs of $13,516 and $14,911 respectively, i.e. there is an incremental cost of $1,395 for MI THR versus STD THR.

MI THR produces 9.53 QALYs, versus STD THR which produces 9.52 QALYs. Therefore, the minimal invasive THR provides a minimal improvement of 0.01 QALYs.

The incremental cost per QALY gained for the MI technique versus the standard technique is $157,156. That would mean that MI technique would be considered cost-effective only if a payer is willing to pay more than $157,156 per QALY gained, which is much higher than the common threshold of around $50,000 per QALY gained and therefore would not be considered a cost effective use of scarce resources.

Table 7: Deterministic Base Case Analysis Results

	MI THR	STD THR	Incremental
Cost	$14,911.70	$13,516.94	$1,394.76
QALY	9.53	9.52	0.01
ICUR			**$ 157,156**

2.5 Probabilistic Sensitivity Analysis

The base analysis in the model was conducted in a deterministic form, such that point estimates for each parameter in the model were entered and an estimate of cost-effectiveness of MI THR and standard THR was calculated. In addition, a probabilistic analysis was conducted using Monte Carlo simulation (MCS) with 5,000 runs, i.e. 5,000 estimates of the costs and QALYs associated with each procedure option were obtained by randomly sampling from each uncertain parameter's probability distribution. The MCS was conducted using the random number generator feature of MS Excel to generate random draws from each input parameters probability distribution. Therefore MCS provided estimates of the uncertainty around the cost-effectiveness ratio.

As proposed, each of the probability distributions used in the original model were carefully examined and decision was made on preserving the one used in the model, substituting with a new distribution or combining previous estimates with new data by respecification.

Input parameters and their associated probability distributions are detailed in Table 8.

The probabilistic sensitivity analysis estimated an incremental cost of $1,399 for MI THR in comparison with standard THR and incremental QALYs of 0.01. Based on the MCS, the incremental cost per QALY gained was $164,307. (Table 9)

Table 8: Probability Distributions of the Input Parameters in the HE model

Parameter Description	Parameter Name	Mean	Distribution
Probability of First Revision	*P_1strev*	0.0113	Beta(890, 78534)
Probability of Second Revision	*P_2ndrev*	0.0275	Beta(141, 5137)
Operative Mortality	*P_omort*	0.0001	Beta(4, 30714)
Probability of DVT	*P_dvt*	0.0075	Beta(12, 1595)
Probability of PE	*P_pe*	0.0025	Beta(4, 1595)
Probability of infection	*P_inf*	0.00004	Fixed
Cost for re-operation due to dislocation	*C_reopdisl*	$7,776	Gamma(23, 335.06)
Cost of DVT for those patients readmitted as inpatient	*C_dvt*	$6,600	Gamma (241.88, 27.29)
Cost of PE	*C_pe*	$7,442	Gamma(179.60, 41.44)
Cost of re-operation due to infection	*C_reopinf*	$24,377	Gamma(155.50, 156.76)
Cost of Std THR operation and first year follow-up Cost	*C_std*	$10,766	Gamma(38650, 27.85)
Cost of MI THR operation and first year follow up	*C_mi*	$12,161	Gamma(322.62, 37.69)

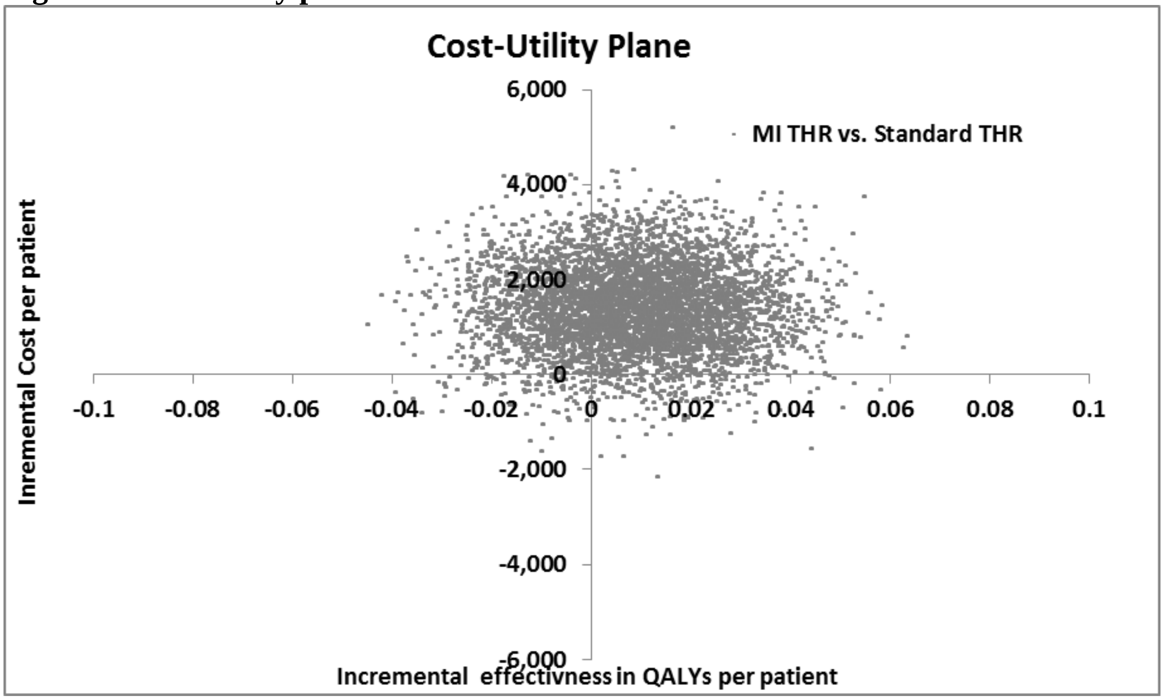
Parameter Description	Parameter Name	Mean	Distribution
Cost of follow up for successful standard, MI or revision THR	*C_followuo*	$45	Fixed
Cost of revision	*C_rev*	$16,199	Gamma (217.57, 74,45)
Pre-operative Utility	*U_preop*	0.6527	Beta(116, 62)
3 months Std THR utility	*U_3mth*	0.7632	Beta(233, 72)
6 months Std THR utility	*U_6mth*	0.8014	Beta (255, 63)
1 year Std THR utility	*U_12mth*	0.8139	Beta (154, 35)
Failure of THR or complications	*U_fail*	0.3200	Beta(309, 656)
Successful Revision	*U_succ*	0.6200	Beta(648, 397)
3 months Incremental MI THR utility	*UI_3mth*	0.0450	Normal(0.045, 0.31)
6 months Incremental MI THR utility	*UI_6mth*	0.0010	Normal(0.001, 0.03)
1 year Incremental MI THR utility	*UI_12mth*	-0.011	Normal(-0.011, 0.033)
Relative Risk of DVT	*RR_dvt*	1	Lognormal(0, 0.0071)
Relative Risk of infection	*RR_inf*	1	Lognormal(0, 0.0071)
Relative Risk of PE	*RR_pe*	1	Lognormal(0, 0.0071)
Relative Risk of revision	*RR_rev*	1	Lognormal(0, 0.0071)
Relative Risk of operative mortality	*RR_omort*	1	Lognormal(0, 0.0071)

Table 9: Probabilistic Base Case Analysis Results

	STD THR	MI THR	Incremental
Cost	$ 13,503.05	$ 14,902.58	$1,399.52
QALY	9.52	9.53	0.01
ICUR			**$ 164,307**

The uncertainty in the cost utility estimate is shown by presenting the joint distribution of costs and effects by means of a scatterplot on the cost-effectiveness plane (Figure 3).

Figure 3: Cost-utility plane for MI THR versus standard THR



The plane represents the four potential results of an economic evaluation with the x axis representing incremental effects (QALYs) and the y axis representing incremental costs. The east half of the plane is associated with an effective treatment, whilst the north half is associated with a more expensive treatment alternative. The results of an evaluation can be placed in each of the four quadrants of the plane and can be characterized in terms of cartography (NW, NE, SE, SW) Scatter plots illustrate the degree of dispersion of the results based on the MCS and the probability that a specific treatment option could be placed in any of the four quadrants. Therefore, scatter plots are an empirical representation of the joint distribution of incremental costs and benefits.

There is a wide scatter in the distribution of the results, reflecting the uncertainty associated with the input parameter estimates. 66.7% of replications are in the NE quadrant of the plane representing an incremental cost and an incremental gain in QALYs associated with MI THR; 3.6% of replications are in the SE quadrant representing cost savings alongside a gain in QALYs; 1.9 % of replications are in the SW quadrant representing cost savings with a loss in QALYs; and, 27.8% of replications are in the NW quadrant representing an incremental cost with an associated loss of QALYs.

The replications fall in all four quadrants and therefore a credible interval for the ICER can not be given. A 95% credible interval is defined by the 2.5th and 97.5th percentile of the distribution of replications. Namely, for the replications that indicate MI THR being either dominated (more costly, less effective) by STD THR (SE quadrant) or to dominate (less costly, more effective) STD THR (NW quadrant), attaching numerical value of ICUR and further estimation of credible intervals is pointless. Furthermore, the replications that fall in the NE and SW quadrant are associated with ratios of costs and QALYs that follow opposite orderings and are not comparable, and therefore combining them in a single credible interval is impossible. The benefits and incremental costs on the x and y axes with the frequency of combinations of costs and effects on the z axis. This requires grouping categorically to allow determination of frequency counts. Figure 4 is an iso-probability contour plot is two dimensional but similar to the three dimensional histogram that supplements the cost utility plane by presenting the frequency of the joint distribution of incremental costs and benefits of 5,000 samples.

Figure 4: Iso Probability Contour Plot of Incremental Costs and QALYs of MI THR versus STD THR

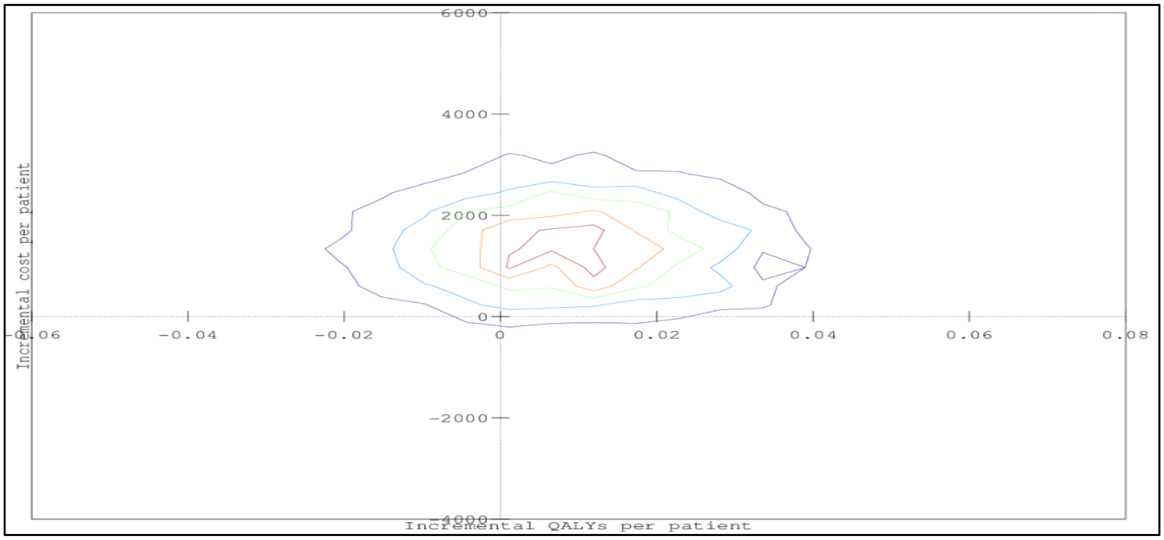
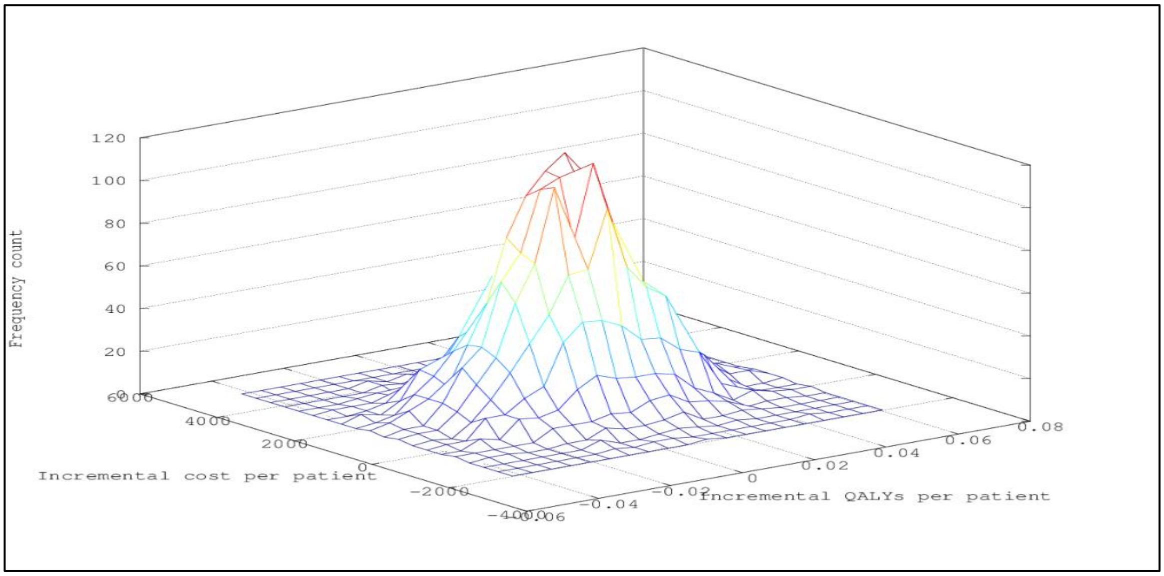


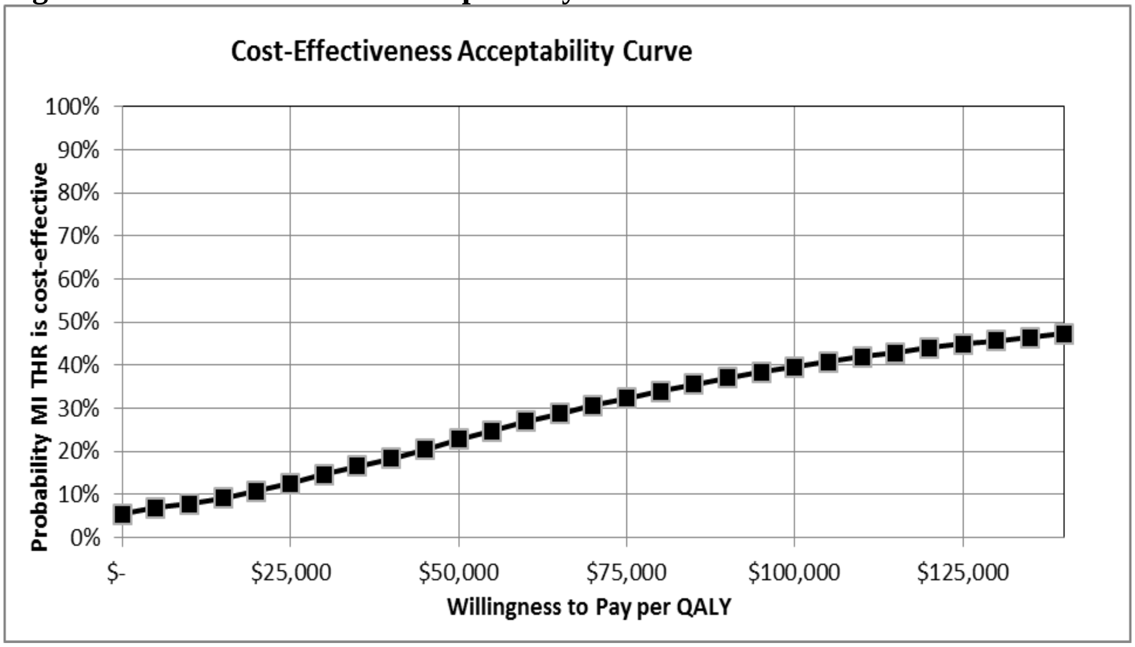
Figure 5: 3-dimensional Histogram of Incremental Costs and QALYs of MI THR versus STD THR



In Figure 6 the probability of cost-effectiveness for the different procedures at different willingness-to pay (WTP) ratios are presented in a form of cost-effectiveness acceptability curve (CEAC).   A CEAC provides a graphical representation of the percentage of replications from the Monte Carlo simulation where the net benefit (NB) of a particular therapy is optimal, i.e. has the greatest NB of all treatment options given a range of values for λ.

Although there is not an explicit value of λ for a cost-effectiveness threshold, as suggested previously, $50,000 is often cited in Canadian cost-effectiveness research as the accepted ceiling for fundable health services. For a WTP $50,000 per QALY there was 20% probability that MI THR is cost-effective compared to standard THR. This decreased to 38% for a WTP of $100,000.

Figure 6: Cost-effectiveness Acceptability Curve for MI THR versus Standard THR



2.6. Comparison of results with the original model

It is interesting to compare, the results of the updated model with the results produced by the original model by Coyle (Coyle et al. 2008). The comparison showed that the ICUR of the base case deterministic scenario is similar to the one produced by the original model (Table 10).

Table 10: Comparison of Deterministic Results among Coyle's Original Model and the Updated Model

	Deterministic Results		Probabilistic Results	
	Original model	Updated model	Original model	Updated model
Incremental Cost	$1,300	$1,395	$1,316	$1,400
Incremental QALY	0.01	0.01	0.008	0.01

	Deterministic Results		Probabilistic Results	
	Original model	**Updated model**	**Original model**	**Updated model**
ICUR	**$ 148,000**	**$164,000**	**$ 148,300**	**$164,307**
Probability MI THR is cost effective, when WTP =$50,000			**47%**	**20%**

The similarities in the results are despite several differences between the analyses. To summarize, there were several changes that were made from the original model:

Model structure. The structure of the model has been changed, so that the updated version of the model assumes no option of non-operative management due to THR failure. Also, subsequent revision THR was differentiated as a separate health state from the first revision THR.

Revision rates. In the original model (Coyle et al 2008), the annual rate of revision was increasing from one year to the next until year 24, derived from the Swedish registry. In the new model that is being used, the annual risk of first revision was fixed and equal to 0.01134, while the annual risk of subsequent revision, if the first one failed, was equal to 0.02751. These probabilities were derived from a more recent source (Lie, 2004), suggesting that the annual risk of subsequent revisions stays constant over time.

Risk of complications. In the original model, four complications were considered: deep venous thrombosis (DVT), pulmonary embolism (PE), dislocation, and deep infection. With the revised model, infection and dislocation were assumed to be a combined complication based on the available data.

Costs. All costs have been updated to reflect 2010 costs, by either inflating the costs of the original mode or by using new data sources.

The results of the comparison showed that the incremental QALYs of MI THR versus STD THR is the same as per both models, but the incremental cost of MI THR versus STD THR is slightly higher in the updated model, which resulted in slightly higher ICUR of STD THR versus MI THR in the updated model when compared to the original model.

Chapter 3

Measuring Parameters' Importance

3.1 Introduction

Uncertainty in parameters is present in many risk assessments and decision analyses and necessarily leads to uncertainty in model predictions. Therefore, an analysis of the degree of uncertainty around the model inputs is often needed.

Sensitivity analysis involves the determination of the change in the response of a model to changes in individual model parameters and specifications (Iman and Helton 1988). Sensitivity analysis includes various analytic methods and is used to identify the main contributors to the variation or imprecision in the outcome. Furthermore, sensitivity analysis can play an important role in model verification and validation and also can be used to provide insight into the robustness of model results when making decisions. Sensitivity analysis can be used to help develop a "comfort level" with a particular model. If the model response is reasonable from an intuitive or theoretical perspective, then the model users may have some comfort with the qualitative behavior of the model even if the quantitative precision or accuracy is unknown (Frey and Patil 2002).

On the other hand, importance analysis involves the use of quantitative methods aiming at identifying the contribution of uncertain input model parameters to output uncertainty (Borganovo 2006). Such methods are called uncertainty importance measures, and allow the ranking of the model input variables by their impact to the model uncertainty. This information helps both in identifying the most important contributors to output uncertainty, and in finding the most effective ways of reducing that uncertainty, by providing information on which parameters it may be worthwhile obtaining more information (Smith and Croke 2007).

Importance analysis has applications in models where any uncertainty has to be eliminated, such as nuclear safety models. In these cases importance analysis will contribute to reducing the uncertainty in the model by identifying those parameters which can lead to system failure. On the other hand, importance analysis has applications in models where some degree of uncertainty can be tolerated and the purpose of importance analysis is to identify which parameters are important in terms of the degree of uncertainty they cause.

Thus, the role of importance analysis is to identify parameters for which it might be worthwhile obtaining more information. This is clearly appropriate for health economic models. Importance analysis has been applied to a range of disciplines such as health economic models (Coyle et al 2003), food safety (Frey and Patil 2002) and gully erosion models (Smith and Croke 2007).

Importance analysis is a basic part of risk assessment that focuses on the uncertainties in the assessment and provides more information on which parameters it is worth obtaining more information. This is of great importance for decision makers as it gives advice on what further information needs to be gathered.

Probabilistic methods such as Monte Carlo simulations have been identified as a suitable base for importance analysis, in that they allow for estimation of likelihood of various model outputs based on a number of sets of input parameters generated by sampling from their probability density distribution (Coyle et al 2003).

3.2. Objective

A number of importance measures have been suggested in the literature. Based on the importance techniques, measures have been classified as either variance based measures, elasticity-based measures, probability-based measures, entropy, or information-value based measure. (Coyle et al 2003) In addition, a literature search was performed to identify additional techniques emerging after Coyle's study in 2003. Three additional techniques, ANCOVA, sequential bifurcation and dominance measure technique were identified. In the following chapter the theory behind each of the techniques assessed by Coyle (Coyle et al 2003), as well as the three additional techniques will be outlined, and an algorithm for their application will be provided. The methods were selected based upon a judgment that they have practical application and are of potential relevance.

3.3 Methods Overview

3.3.1 Expected Value of Partial Perfect Information

Expected value of perfect information (EVPI) is an information-based measure of the reduction in opportunity loss associated with obtaining perfect information (no uncertainty) on a parameter and can be seen as a measure of decision sensitivity (Coyle 2004). The decision sensitivity is determined by the probability that a decision based on existing information will be wrong and the cost consequences if the wrong decision is made, since perfect information could eliminate the possibility of making wrong decision (Claxton1999).

In mathematical terms, EVPI is defined as:

$$EVPI = \int_{-\infty}^{0} f(INB) INB \, dINB$$

where INB = Incremental Net Benefit.

The interpretation of the mathematical formula is that EVPI is a product of the probability of a change in what is the optimal treatment and the average change in INB as a result of such a change. As Felli and Hazen (1998) stated EVPI methodologically represents a natural extension of probabilistic sensitivity analysis by focusing simultaneously on the probability of a decision change and the change in payoff corresponding with such a decision change (Felli and Hazen 1998). The higher the EVPI, the larger the opportunity cost of an incorrect decision viewed at the point at which the uncertain decision is being made (i.e., the more costly is the uncertainty) (Briggs et al 2012). EVPI is reported in monetary terms, i.e. using net monetary benefit (NB). Because the net monetary benefit depends on the ICER threshold, EVPI is reported for a specified ICER threshold (Briggs et al 2012).

Based on whether uncertainty around a single input parameter or all input parameters is under consideration, EVPI can be defined as global EVPI or partial EVPI (most commonly known as expected value of partial perfect information EVPPI). For the purpose of this thesis, i.e. measures of importance of single variables, EVPPIs will be the focus.

Using the Bayesian approach, EVPPI can be expressed as:

$$EVPPI(x_i) = E_{x_i}(\mathrm{max}_{\mathrm{t}}[E_x\,(NB_t)] \mid x_i) - max_t[E_x(NB_t)]$$

The challenging term in the above equation is $E_{x_i}(\mathrm{max}_{\mathrm{t}}[E_x\,(NB_t)] \mid x_i)$, which involves a sequence of expectation, maximization and expectation. The inner expectation evaluates the net benefit over the remaining uncertain parameters conditional on x_i. The outer evaluates the net benefit over the parameters of interest x_i.. EVPPI cannot be solved in closed form, and therefore it requires estimation. It is important that the two expectations are evaluated separately because of the need to compute a maximum between the two expectations, which contributes to computational complexity of EVPPI (Brennan et al 2007).

EVPPI has a number of advantages: (i) it combines both the importance of the parameter (how strongly it is related to differences in NB) and its uncertainty, (ii) it is directly related to whether the uncertainty matters (whether the decision changes for different possible values); (iii) does not require a linear relationship between inputs and outputs (Claxton 2008).

In layman's term, EVPPI can be interpreted as the expected benefit by completely resolving uncertainty around an individual input parameter.

In the next section, three different methods for calculating EVPPIs will be outlined.

3.3.1.1. EVPPI - Double MCS Method

Brennan, Chilcott and colleagues amongst others have suggested a method of calculating EVPPIs which involves solving both the inner and outer integration through a two stage Monte Carlo simulation (Coyle et al. 2003).

If MCS with N_I replication is conducted for the inner integration, and N_O replication for the outer integration, than the formula for calculating EVPPI can be represented as summations as follows:

$$EVPPI(x_i) = \frac{1}{N_O}\sum_{1}^{N_O}(max_t(\frac{1}{N_I}\sum_{1}^{N_I}[(NB_t) \mid x_i])) - NB_{t*}$$

where NB_{t*} is the net benefit of the optimum therapy t^* from the base analysis.

Algorithm 1: EVPPI with Double-MCS Method

1. For the parameter of interest, sample number of random values from its probability density functions. (outer MCS)
2. For each value of the parameter chosen in step 1, conduct MCS by sampling from the probability density functions of all other parameters (Xc) (inner MCS).
3. For each simulation conducted in step 2, subtract the expected value of the net benefit of the optimum therapy from the maximum of the expected values of the net benefit over all therapeutic options.
4. Repeat steps 2.to 3 for all values selected for the parameter of interest in step

Two- stage MCS method is not resistant to Monte Carlo biased maxima. Although the estimates of the expectations for each decision are themselves unbiased, the resulting estimate of the maximum over all the decisions will always be biased upwards (Brennan et al 2007). This is because, if X and Y are random quantities, then in general max(E[X],E[Y]) is upward biased if E[X] and E[Y] are estimated with error.

Typically for MCS larger sample sizes will increase the accuracy of computation i.e. reduce the likely error (variance). In considering the number of samples required for outer and inner MCS the precision required of the resulting partial EVPI estimate need to be considered (Brennan et al 2007). An equal number for inner and outer sampling is not necessary and may not be efficient. Brennan concluded that improving the accuracy of partial EVPI estimates requires proportionately greater effort on the inner level than the outer level (Brennan et al 2007).

3.3.1.2 EVPPI- Quadrature Method

The quadrature method for estimating EVPPI has been introduced by Coyle, suggesting that the outer integration across the probability density functions of x_i can be achieved through numeric quadrature approximation using Simpson's rule (Coyle et al 2001, Coyle et al 2003).

Simpson's rule is a numerical approximation of the area under a curve through a series of parabolic segments.

Let $f(x)$ be a function over an interval $[a,b]$. Suppose the interval is divided into $2m$ subintervals $[x_{k-1}, x_k]$, k=1,..2m, with equal length $h = \frac{b-a}{2m}$.

Then

$$\int_a^b f(x) = \frac{h}{3}(f(a) + f(b)) + \frac{2h}{3}\sum_{k=1}^{m-1} f(x_{2k}) + \frac{4h}{3}\sum_{k=1}^{m} f(x_{2k-1})$$

Using Simpson's rule, the calculation of EVPPI can be much more efficient, such that the outer expectation can be calculated implementing the Simpson's rule.

Algorithm 2: Quadrature Method for Calculating EVPPI

1. A set of values {xi| i=1...N} is determined for the parameter of interest. Select the values should such that $x_{i+1} - x_i = d$, for for $\forall i = 1 \ldots N, d = \overline{const}$ and the probability density function close to 1.
2. For each value of the parameter chosen in step 1, conduct a MCS by sampling from the probability density functions of all other parameters (Xc).
3. For each simulation conducted in step 2, subtract the net benefit of the optimum therapy from the base analysis from the maximum net benefit over all therapeutic options.
4. Weight each estimate from step 3 by the probability density for the specific value of the parameter.
5. Estimate EVPPI by integrating across the probability density function using Simpson's rule.

3.3.1.3 Single MCS Method

Felli and Hazen suggested another method of calculating EVPPI using a single MCS (Felli and Hazen 1998). They suggested a procedure that can be applied successfully "when all parameters are assumed probabilistically independent and the pay-off function is multi-linear i.e. linear in each individual parameter" (Brennan et al 2007). Using this method, the outer level expectation over the parameter set of interest is as per the double-stage MCS method, but the inner expectation is replaced with net benefit calculated given the remaining uncertain parameters X^C set at their prior mean, i.e.

$$EVPPI(x_i) = E_{x_i}(\mathrm{max_t}[\,(NB_t)] \mid X_i{}^C) - max_t[E_x(NB_t)]$$

In order this method to be valid the following equations need to be valid:

$$E_{x_i}(\mathrm{max_t}[\,(NB_t)] \mid X_i{}^C) = E_{x_i}(\mathrm{max_t}[E_x\,(NB_t)] \mid x_i)$$

This equation will be true if and only if:

1. For each treatment ***t,*** NB_t can be expressed as a sum of products of components of X^C.

2. The components of X^C are mutually probabilistically independent.

Thus, this approach is generally appropriate for analyses using decision tress but not for Markov models.

To summarize, this approach performs a single MCS, allowing parameters of interest to vary, keeping remaining uncertain parameters constant at their prior means. It is much more efficient than the double MCS method, since many model runs are replaced by a single run, however its application is limited to only simplistic models such a decision trees.

Algorithm 3: EVPPI - Single MCS method

1. Conduct a MCS by sampling from the probability density functions of the parameters of interest (Xi) with all other parameters fixed at their expected value (Xc= E(Xc)).
2. For each replication within the MCS calculate the difference between the net benefits of the optimal treatment as previously identified and the maximum net benefits across all treatments.
3. EVPPI is the expected value from step 2.

3.3.2 Correlation Coefficients

Standard non-parametric measure of statistical dependence between two variables is the correlation coefficient, either Pearson's correlation coefficient or Spearman rank correlation.

Pearson's correlation coefficient (R_p) between two variables is defined as the covariance of the two variables divided by the product of their standard deviations and is extensively used to measure the strength of linear dependence between two variables. Spearman rank correlation (R_s) is simply the Pearson correlation coefficient calculated on rank transformed data. With the rank transformation, data are replaced with their corresponding ranks, and then the usual correlation procedure is performed on these ranks.

When there is non-linear dependence between the variables, it has been shown that Spearman rank correlation (R_s) is more appropriate than Pearson correlation (Hofer 1999).

$$R_S = \frac{\sigma_{r_{x_i}, r_{INB}}}{\sigma_{r_{x_i}} \sigma_{r_{INB}}},$$

where

r_{x_i}= rank transformed x_i

r_{INB}= rank transformed *INB*

Correlation analysis measures the effect of one input at a time (i.e., x_i) on the output (y) (Helton and Davis 2002). It assesses how well the relationship between two variables can be described using a monotonic function.

To summarize, Spearman's correlation coefficient is normalized version of the covariance of the ranks of the inputs. Since covariance is a measure of how much two random variables change together, normalized version of covariance help in quantifying that change. Therefore in layman's terms, Spearman correlation measures the degree to which the rank orderings for the input variable and the output variable change together.

It has been argued that neither Pearson nor Spearman coefficients can capture complex dependencies such as thresholds or directly deal with interactions among inputs (Mokhtari 2006).

Algorithm 4: Correlation Coefficients

1. Conduct MCS with high number of replications by sampling from the probability density functions of all parameters (x_i).
2. Calculate the rank of the values for all parameters (x_i) and the output (INB).
3. Calculate the correlation Rs between the rank of the input parameters (r_{xi}) and the rank of the output (r_{INB}), calculated as covariance of rxi and r_{INB} $\sigma_{r_{x_i},r_{INB}}$, divided by the product of their standard deviations, $\sigma_{r_{x_i}} * \sigma_{r_{INB}}$.

3.3.3 Regression Coefficients

Regression coefficients have also been suggested as a possible measure of parameter importance (Iman and Helton 1988, Helton 1993).The generalized form of a simple linear regression equation is:

$$INB = \alpha + \sum \beta_i x_i$$

where

x_i = input parameters,

α, β_i = regression coefficients.

Regression techniques allow measuring importance of the parameters based on regression coefficients. However, since these values are function of the relative magnitude of the parameters, the regression coefficients needs to be standardised. The effect of the standardization is to remove the influence of units and treat all parameters on an equal basis.

There are several methods for standardizing regression coefficients suggested in the literature, the most common being standardization through estimation of standardized regression coefficients (SRC) or rank regression coefficients (RRC). RRCs are the regression coefficients (β_i) obtained from a regression analysis based on rank transformed data (Saltelli et al. 2000). SRCs are the coefficients from a regression analysis weighted by the ratio of the standard deviations of the input parameter and the output (Hamby 1994, Saltelli et al. 2000).

SRC can be calculated as follows:

$$SRC(x_i) = \beta_i \frac{\sigma_{x_i}}{\sigma_{INB}}$$

Where β_i is obtained from the linear regression $INB = \alpha + \sum \beta_i x_i$

Another method of standardizing regression coefficients is conducting regression analysis on already standardized parameters.

For the purpose of this research, the aforementioned method of standardizing regression coefficients by weighting by the ratio of the standard deviations of the input parameter and the output was used.

In layman's term, the regression coefficient tells us an average change in output (*INB*) for a unit change in input parameter (x_i), given that other parameters do not vary (Bring 1994).

Algorithm 5: Standardized Regression Coefficients

1. Conduct MCS with high number of replications by sampling from the probability density functions of all parameters (x_i)
2. Run regression analysis with x_i being independent parameters and INB dependent variable
3. Standardize the regression coefficients βi obtained in step 2 by weighting it with standard deviations of x_i and INB.

3.3.3 Elasticity Coefficient

Elasticity is a common measure in economics of the change in the outcome as a result of a change in the input parameter. In mathematical terms, point elasticity is defined as elasticity at a single input value:

$$\varepsilon_{x_i} = \frac{dINB}{dx_i}\frac{\mu_{x_i}}{\mu_{INB}}$$

Two elasticity methods have been used as importance analysis measures, actual elasticity coefficient (AEC) and absolute relative overall sensitivity (AROS).

AEC is a calculated as a product of point elasticity of a single parameter and its coefficient of variation.

Both point elasticity (ε) and AEC can be calculated without MCS, however they both have some serious disadvantages. Point elasticity does not incorporate the uncertainty levels of the input variable, and it does not reflect the variability of elasticity across different range of input variables (Steen and Erikstad 1996). AEC does not encounter for the variability in elasticity across a range of input parameters.

Absolute relative overall sensitivity (AROS) is an alternative measure that accounts for both uncertainty in the input variable and the variability over the range of input parameters. AROS estimates the responsiveness of outcomes to the values of the input parameter by linear regression. The regression coefficient is then used to estimate the overall elasticity across range of values. This method requires assumption of linearity between input parameters and the output (Coyle 2004).

Coyle proposed an alternative method that does not require the assumption of linearity. He defined the elasticity coefficient (EC) as an expectation of elasticity over the value of the input parameter. Therefore EC is calculated for each input parameter by integration of elasticity values over all possible values of the parameter weighted by its probability density function. The integral can be calculated through MSC and numerical approximation using Simpson's rule.

To summarize, elasticity is a mathematical tool which can be used to seek the strength of the relation between any two variables. In layman's terms, the elasticity coefficient gives the ratio of the percent change in INB to percent change in input parameter, i.e. the sensitivity of INB to changes in input variables. Elasticity analysis ranks the importance of parameters based on their potential to influence INB when changed by a set percentage from the mean.

Algorithm 6: Elasticity Coefficient

1. A set of values {xi| i=1...N} is determined for the parameter of interest. Select the values such that $x_{i+1} - x_i = d$, for for $\forall i = 1...N, d = \overline{const}$ and the probability density function close to 1.
2. For each value of the parameter determined in step 1. Estimate INB by conducting MCS with sampling from the probability density distributions of all other parameters (X_C)
3. For each set of values from set 2, estimate point elasticity (ε_{x_i})
4. Weight each ε_{x_i} from step 3 by the probability density function of x_i.
5. EC is calculated by integrating across probability density function of the parameter of interest using Simpson's rule

3.3.4 Contribution to Variance Measure

Decomposing the total output variance and ranking the input parameters in a model based on their partial contribution to the total variance of an output has been frequently used as an importance measure (Coyle et al 2003). Contribution to variance (CV) can be simply defined as the proportion of the variance of the outcome explained by the input parameter (Coyle et al 2003).

$$CV = \frac{\sigma_{INB}^2 - \sigma_{INB|x_i}^2}{\sigma_{INB}^2}$$

INB variance is estimated by MCS with sampling through the probability distributions of all input variables.

CV involves ranking the parameters according to how much the unconditional variance of the output is reduced by fixing the input parameters to their true values (Saltelli et al 2002). Since the true value of the input variables is unknown, it is sensible to look at all possible values of the input parameters.

The advantage of this method is that this measure is valid, even if the model is nonadditive and the parameters are correlated (Satelli and Tarantola 2002).

Algorithm 7: Contribution to Variance

1. A set of values {xi, i=1...N} is determined for the parameter of interest. Select the values such that $x_{i+1} - x_i = d$, for for $\forall i = 1 \dots N$, $d = \overline{\mathrm{const}}$ and the probability density function close to 1.
2. For each value of the parameter selected in step 1, estimate the variance of INB by sampling through the probability distributions of all other input variables.
3. Weight each estimate from step 3 by the probability density function for the specific value.
4. Estimate the variance of INB conditional to xi is by integrating across probability density function using Simpson’s rule.

3.3.5 Generalized Sensitivity Analysis (Maximum Separation Distance)

The generalized sensitivity analysis is a probability-based measure that involves comparing two cumulative distribution functions for each parameter of interest. The cumulative functions are derived by partitioning the sample for the input parameter based on a dichotomy relating to outcome. This can be done by categorizing by INB greater than 0 and INB less than 0.

Parameter importance can be measured by the maximum vertical distance between the two cumulative functions (maximum separation distance (MSD)) defined as (Coyle 2004):

$$MSD(x_i) = max_k[p(\, x_i < k \mid INB < 0\,) - p(\, x_i < k \mid INB \geq 0\,)]$$

If there is a large separation distance between the two cumulative distribution functions, then the output is sensitive to the input variable. This measure can be measured with Kolmogorov-Smirnov statistic (K-S). In fact K-S statistic provides the probability of MSD between the two cumulative distribution functions that could have occurred if the two distributions came from a same distribution. The bigger the difference between the two probability density functions, the larger the impact of the parameter to the INB.

Therefore, the most important parameters can be determined by ranking the input parameters by MSD or K-S statistic.

Algorithm 8: Maximum Separation Distance

1. Conduct MCS by sampling through the probability distributions of all parameters of interest.
2. Create 2 separate data sets A and B, out of the MCS run, such that
3. $A = \{ r_i \mid INB_i > 0 \}$ and $B = \{ r_i \mid INB_i \leq 0 \}$
4. Where $r_i =$ i-th run
5. For each value (k) within the range of values of the input parameter xi obtained from the MCS, calculate the proportion of A where xi is less than the value (the cumulative density).
6. For each value (k) within the range of values for each input parameter xi obtained from the MCS, calculate the proportion of B where xi is less than the value (the cumulative density)
7. For each value within the range of values of xi, calculate the absolute difference between the cumulative densities for A and B.
8. Calculate the maximum value from step 5, which is MSD.

3.3.6 Entropy Based Measures

In information theory, entropy is defined as a measure of the uncertainty associated with a random variable (Ihara 1993). In this context, the term usually refers to the Shannon entropy, which quantifies the expected value of the information. Entropy is the negative of the expectation over a parameter of the product of the probability density of the parameter and the log of the probability density.

$$H(INB) = -E_{INB}\left[f(INB).\log(f(INB))\right]$$

Intuitively, the entropy gives a measure of the uncertainty around the random variable. It is sometimes called the missing information: the larger the entropy, the less a priori information one has on the value of the random variable. This measure is roughly speaking the logarithm of the number of typical values that the variable can take (Info Phys Comp 2007).

In information theory, the mutual information (MI) of two random variables or relative entropy is a quantity that measures the mutual dependence of the two random variables.

Formally, the mutual information of two continuous random variables X and Y can be defined as:

$$I(X;Y) = \int_Y^{\cdot} \int_X^{\cdot} p(x,y) \log\left(\frac{p(x,y)}{p(x)p(y)}\right) dxdy,$$

where $p(x,y)$ is now the joint probability density function of X and Y, and $p(x)$ and $p(y)$ are the marginal probability density functions of X and Y respectively.

Mutual information can be also expressed as

$$I(X;Y) = H(X) - H(X \mid Y) = H(Y) - H(Y \mid X)$$

Applying the above formula in the context of the health economic evaluation, MI can be defined as the difference between the entropy of INB as estimated above and the entropy of INB conditional on x_i ($H(INB|x_i)$) (Coyle 2004).

$$I(INB;x_i) = H(INB) - H(INB \mid x_i),$$

where:

$$H(INB \mid x_i) = E_{x_i}[-\mathrm{E_{INB}}[f\big(INB \mid x_\mathrm{i} log\, g\big(\mathrm{f}(\mathrm{INB} \mid \mathrm{x_i}\,)\big)]\,]$$

In layman's terms, mutual information measures the information that two parameters share. MI measures how much knowing one of these parameters reduces uncertainty about the other. If entropy H(X) is regarded as a measure of uncertainty about a random variable, then H(X|Y) is a measure of uncertainty remaining about X if Y is known. Therefore,

$I(X;Y) = H(X) - H(X \mid Y)$ can be interpreted as the amount of uncertainty in X minus the amount of uncertainty in X which remains if Y is known, i.e. the amount of uncertainty in X which is removed by knowing Y.

Algorithm 9: Mutual Information

1. A set of values {xi| i=1...N} is determined for the parameter of interest. Select the values such that $x_{i+1} - x_i = d$, for for $\forall i = 1...N, d = \overline{const}$ and the probability density function close to 1.
2. For each value of the parameter chosen in step 1, conduct MCS by sampling from the probability density functions of all other parameters (Xc).
3. Based on the results of the MCS, a conditional estimate of entropy is estimated as above.
4. Each estimate from step 3 is weighted by the probability density for the specific value of the parameter.
5. $H(INB|x_i)$ is then estimated by integrating across the probability density function using Simpson's rule.

3.3.7 Dominance Measure

The dominance measure methodology is based on the general method presented for comparing the relative importance of parameters in multiple regressions, known as dominance analysis. (Budescu 2003, Azen and Budescu 2003). Dominance analysis is a procedure that is based on an examination of the R^2 values for all possible subset models.

In Budescu's (2003) approach, called dominance analysis (DA), parameter's contribution to prediction was defined as the squared semipartial correlation, and the level of analysis was defined as all $2^{(p-2)}$ subsets for which the comparison of each pair of parameters is relevant, where p is the total number of parameters. In other words, one parameter is more important than another if it would be chosen in all possible subset models where only one parameter of the pair is to be entered.

For the purpose of this research, the proposed approach was modified slightly such that only N different subsets were examined, with N=number of parameters in the model. Each subset was defined such that one parameter, i.e. the parameter of interest was excluded from the set. Further, linear regression analysis was conducted for each subset S_i, i =1, ...,N, as well as for S, where S represents the full set of parameters. We define dominance measure to be

$DM = \frac{R_i^2}{R_S^2}$, where

R_i^2 = coefficient of determination R^2 from the regression analysis on subset S_i,

R_S^2 = coefficient of determination R^2 from the regression analysis on full set S.

The advantage of this method is that it has a natural and intuitive interpretation. Namely, it address the question of whether variable x_i is more or less (or equally) important than variable x_j in predicting *Y*, where the definition of importance is straightforward—x_i is considered more important than x_j if it contributes more to the prediction of the response than does x_j (Azen and Budescu 2003).

Algorithm 10: Dominance Measure

1. Conduct a MCS by sampling through the probability distributions of all parameters of interest
2. Conduct a regression analysis using all input parameters as independent variables and INB as a dependant variable.
3. Conduct a regression analysis using all except the input parameter of interest as independent variables and INB as a dependant variable.
4. Obtain R^2 for the regression analyses from step 2.and 3.
5. The relative ratio of the R^2 is the dominance measure of importance for the parameter of interest.

3.3.8 ANCOVA

Analysis of co-variance known as ANCOVA is another option that can be used to measure parameter importance. ANCOVA combines one-way or two-way analysis of variance with linear regression and can summarize the proportion of the variance in the output due to variation to the input parameters. ANCOVA evaluates whether the population means of the dependent variable, adjusted for differences on the covariate(s), differ across the levels of the independent variables (Tabachnick and Fidell 1996).

ANCOVA can summarize the proportion of the variance in the output variable “explained” by the variation in the input variable (Briggs et al 2006).

The major limitation is that it assumes linear relationship between inputs and output variables, and therefore it is an approximation when dealing with non-linear models.

Algorithm 11: ANCOVA Standardized Coefficients

1. Conduct a MCS by sampling through the probability distributions of all parameter of interest.
2. Conduct ANCOVA using all input parameters as independent variables and INB as a dependant variable.
3. Obtain coefficients from step 2.

3.3.9 Sequential Bifurcation

Sequential bifurcation is a group-screening method for searching important parameter (Bettonvil and Kleijnen 1997). This method which may be thought of as a generalisation of classical binary search (the objective of this latter technique being to find the single most important factor), was developed for deterministic simulations (Cheng 1997).

Group-screening methods can be very useful for models with large numbers of factors. The fundamental idea is to identify the important/unimportant factors as a group, and reduce time.

If a group is considered to be important, then subgroups or individual factors within the group should be further screened; if a group is not considered to be important, then the whole group can be classified as unimportant (Wan et al 2006).

The input/output behaviour of the model can approximated with a first-order polynomial

$$INB = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_N x_N \ ,$$

where N is the number of input variables.

In order to avoid cancelation, the direction of influence that each parameter has on the output has to be known.

This method is based on the deterministic framework, and it does not comfort with standard practices of Bayesian statistics. Also, there is not a standard procedure to determine the threshold B, and therefore the model also involves certain subjectivity in determining when parameter can be classified as important or unimportant. As discussed earlier, some a priori knowledge of the direction of influence of the input variables is also required. Some knowledge of the importance of input parameters is also beneficial to for forming the subgroups, such that the group of parameters that are not important can be eliminated in the earlier stages.

The algorithm for sequential bifurcation is outlined below.

Algorithm 12 Sequential Bifurcation

1. For all model parameters select two values that will generate high and low value for the outcome (Hi and Li), $x_i^1 = x_{0,} x_i^2 = x_0 + \delta$.

$$\text{If } \text{INB}(X \mid x_i = x_i^1) > \text{INB}(X \mid x_i = x_i^2), \text{ then } x_i^1 = H_i, x_i^2 = L_i$$

2. Transform the input variable into binary values, such that

$$x_i = \begin{cases} 0, & x_i = L_i \\ 1, & x_i = H_i \end{cases}$$

3. Let's denote INB (n) the value of the output when input parameters x_1 to x_n are switched on, and the rest are switched off, i.e. x_1,...x_n =0, and x_{n+1}, ...x_N = 1. Calculate INB (0) and INB (N) by setting all parameters to their lower values and their higher values. It would be expected that INB (0) < INB (N), if not, then none of the input parameters are important.

4. Calculate the estimator as $\beta_{1\_N} = \frac{\text{INB}_{(N)} - \text{INB}_{(0)}}{2}$

5. Input parameters are bifurcated (split into 2 subgroups with equal sizes) by

$$X_{1,i} = \{x_1, x_2, \ldots x_i\} \text{ and } X_{i+1,N} = \{x_{i+1}, x_{i+2}, \ldots x_N\}$$

6. INB (i) and consequently β_{1_i} and $\beta_{i+1,N}$ are calculated. If $\beta_{1\_i} > B$, where B is the threshold for determining if parameter importance is achieved or not, then the subset $X_{1,i}$ is further bifurcated into smaller subgroups until the individual parameter that are important are identified. If $\beta_{1\_i} < B$ then it can be concluded that none of the parameters$\{x_1, x_2, \ldots x_i\}$ are important.

Chapter 4

Results

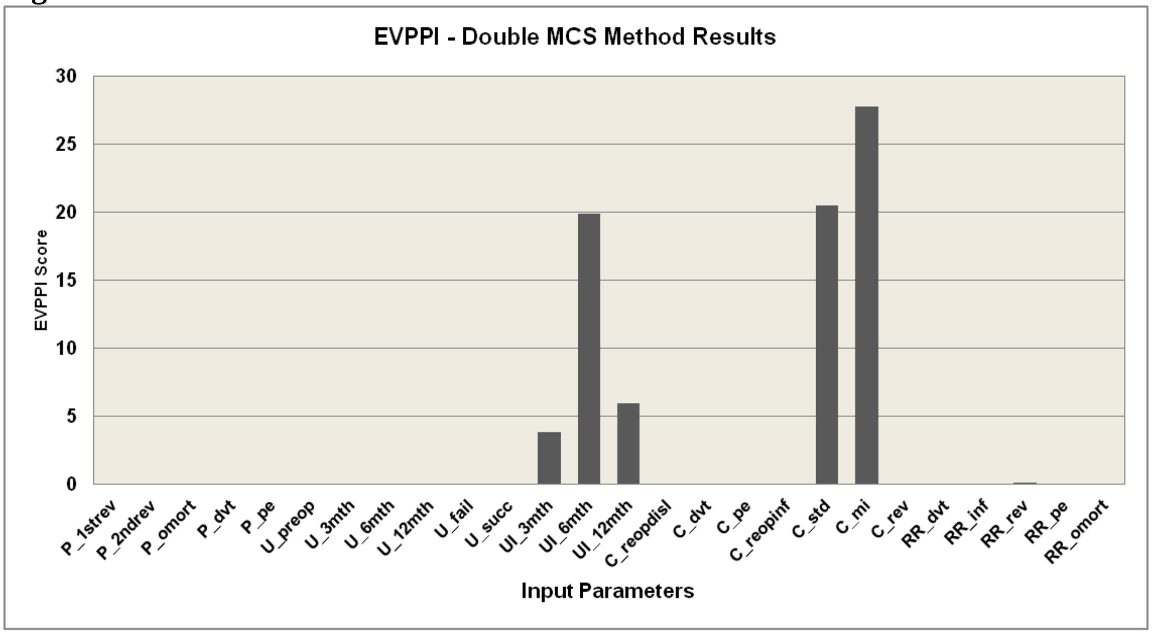
In the following section the results of all methods described in previous section will be presented. A comparison among the results of the different methods will be provided as well as comparison among results of the different methods

4.1 EVPPI Results

4.1.1 EVPPI Double MCS Method

The results of the double MCS method are presented in Figure 7.

Figure 7: EVPPI Double MCS Method Results



Cost of standard THR and cost of MI THR had the highest EVPPI, followed by incremental utility values. The relative risk of revision had little information value, and the other parameters had no information value. For these latter parameters EVPPI is exactly zero, as per the mathematical formula of EVPPI:

$$E_{x_i}(\mathrm{max_t}[E_x\,(NB_t)] \mid x_i) = \,max_t[E_x(NB_t)]$$

This occurs because for any value of the parameter, the NB of STD THR is greater than the NB of MI THR, i.e. the decision that STD THR is an optimal treatment would not change.

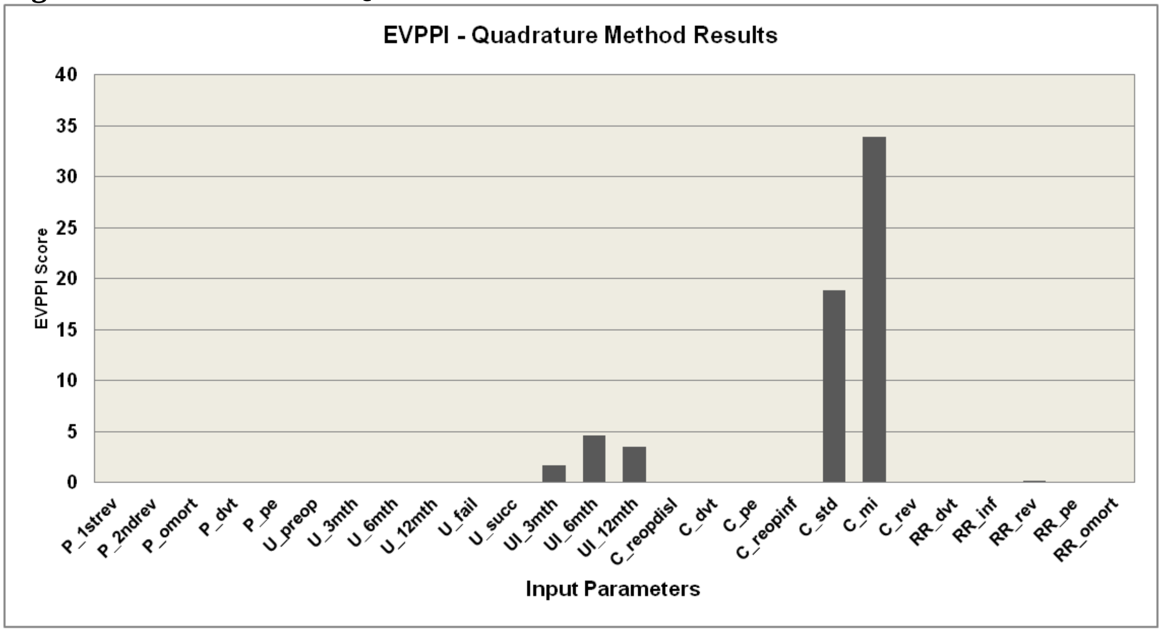
The double MCS method was conducted by using 5000 replication for the inner and 1000 replication for the outer MCS, resulting in a total of 5,000,000 replications. Increasing the number of replications would give results converging more closely to the true values of EVPPI, on expense of increasing the run time. This method was proven to have a long run time if executed in Excel.

Intuitively, it would be expected that the level of uncertainty of making a right decision would peak if λ is equal or close to the ICUR. However, this is not always the case as it is dependent on the relative uncertainty over variables contributing to incremental cost and those contributing to incremental outcome. Coyle et al. explored the relationship among λ and EVPPI in his study and concluded that as the threshold value increased, the importance of utility values increased and the importance of cost parameters as well as of transitional probabilities also decreased (Coyle et al 2003).

4.1.2 EVPPI Quadrature Method

The results of the MCS quadrature method are presented in the Figure 8.

Figure 8: Results EVPPI Quadrature Method



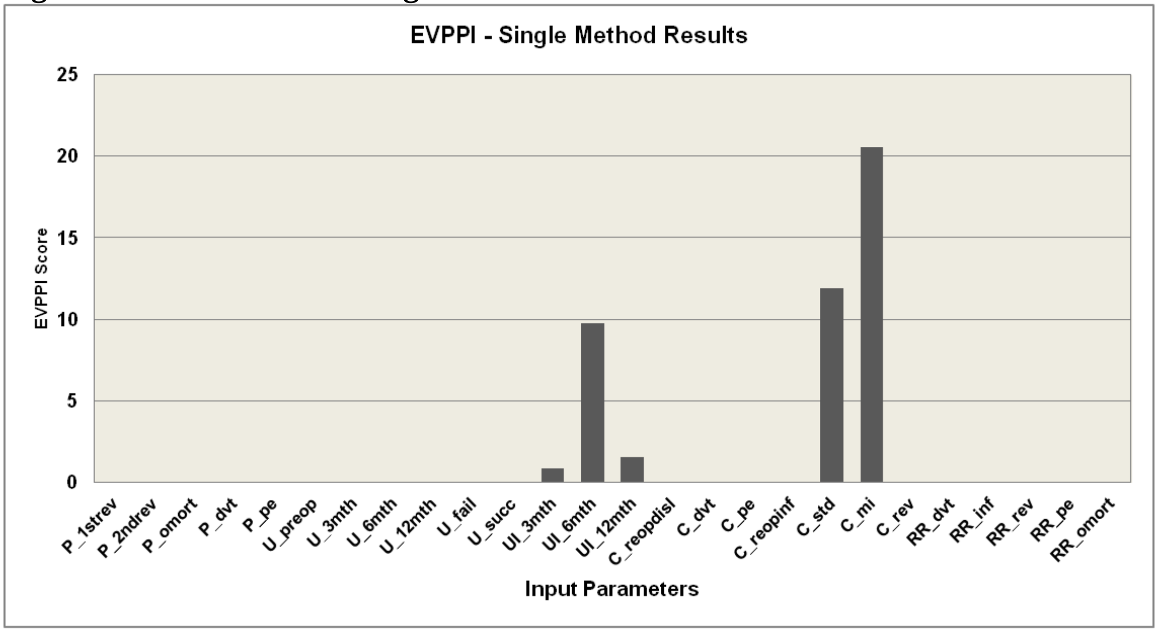
The results from the quadrature method are consistent with the results from the double MCS method in terms of ranking the parameters by their EVPPI values. The costs of standard and MI THR, along with the incremental utilities were similarly identified as parameters with the highest value of information. The quantitative values of EVVPIs slightly differ from the two methods, with EVPPIs derived from quadrature method being slightly lower than those of the double MCS method, which might be due to the number of replications in the double MCS method and/or the extra potential for MCS maxima bias due to the double use of MCS.

The quadrature method was conducted by using 5000 replications for the inner MCs and 101 values in the quadrature estimation of the outer integral, resulting in a total of 505,000 replications. In terms of computational efficiency, the quadrature method has significantly lower run-time than the double MCS due to the lower number of replications required for executing the method. However, this method is methodologically more complex than the double MCS method and requires more sophisticated calculation.\

4.1.3 EVPPI Single Method

The results of the EVPPI single MCS method are summarized in Figure 9. Although this method is appropriate only when all parameters are assumed probabilistically independent and INB is multi-linear in X^c, the method was still applied to all parameters, even if the conditions were not satisfied. In that sense, the single MCS method can be seen as a proxy method for calculating EVPPIs.

Figure 9: EVPPI Results- Single MCS Method



Although in terms of selection of parameters the ranking of parameters with the single MCS method is consistent with the results of the double MCS and quadrature methods, this method resulted with quantitatively slightly different results than the other two methods.

The single MCS method is computationally efficient methods of estimating EVPPI with a very short run-time requiring only one MCS of 5000 replications per parameter.

4.1.4 Comparison of EVPPIs Methods

All implemented EVPPI methods gave consistent results with respect to the identification of parameters with non-zero value of information. The following tables compare the results obtained from the three methods for calculation EVPPIs for all parameters. Table 11 summarises the results in terms of the EVPPI values, while Table 12 contains calibrated values on a scale 0-100 as well as their ranks.

Table 11: Comparison of results for the parameters that resulted with non-zero EVPPI

Parameter^	Double MCS method	Quadrature method	Single MCS method
C_mi	27.76	33.88	20.53
C_std	20.48	18.89	11.93
UI_6mth	19.92	4.60	9.73
UI_12mth	5.96	3.50	1.58
UI_3mth	3.84	1.64	0.84
RR_rev	0.08	0.10	0.00

C_mi=Cost of MI THR; *C_std* = Cost of STD THR; *UI_3mth* = Incremental Utility at 3 months; *UI_6mth* = Incremental Utility at 6 months; *UI_12mth* = Incremental Utility at 12 month; *RR_rev*= Relative Risks of Revision

^Only parameters with non-zero EVPPI are included. The rest of the parameters resulted with zero EVPPI.

Table 12: Comparison of EVPPI Calibrated Scores and Ranks for the parameters that resulted with non-zero EVPPI

Parameter^	Double MCS method	Quadrature method	Single MCS method
C_mi	1 (100)	1 (100)	1 (100)
C_std	2 (73.76)	2 (55.75)	2 (58.11)
UI_6mth	3 (71.75)	3 (13.57)	3 (47.40)
UI_12mth	4 (21.48)	4 (10.33)	4 (7.67)
UI_3mth	5 (13.83)	5 (4.84)	5 (4.11)
RR_rev	6 (0.29)	6 (0.30)	6 (0.00)

C_mi=Cost of MI THR; *C_std* = Cost of STD THR; *UI_3mth* = Incremental Utility at 3 months; *UI_6mth* = Incremental Utility at 6 months; *UI_12mth* = Incremental Utility at 12 month; *RR_rev*= Relative Risks of Revision

^Only parameters with non-zero EVPPI are included. The rest of the parameters resulted with zero EVPPI.

As shown, the results are consistent across the three methods. The main observed difference is that the single MCS method, as opposed to the other two methods, did not identify risk of revision as a parameter with non-zero EVPPI. With all three methods, cost of STD and MI THR had the highest EVPPI, followed by incremental utilities and relative risk for revision. All the other parameters had zero EVPPI based on all three methods. Also, while in general there was no significant difference among the quantitative scores across the three methods, the quadrature method resulted with the EVPPI of incremental utility for 6 months being significantly lower than based on the single and double MCS method. The single MCS method resulted with lower EVPPIs overall, although as noted, the ranking of the parameters was consistent across all techniques.

It should be noted that all methods are subject to Monte Carlo maxima error. The higher the number of replications, the closer the results will be to the true EVPPI value, if an appropriate method is used.

In terms of computational efficiency, the single MCS method has the shortest run-time as it requires the least number of replication (5000 per parameter). The quadrature method was executed with 505, 000 replications which resulted in a longer runtime than the single MCS, but more computationally efficient than the double MCS that required 5,000,000 replications.

4.2 Correlation Coefficients Method

Results of the correlation coefficients method are presented in Figure 10, as well as in Table 13 where the calibrated scores of rank correlations are being compared with the result of the three EVPPI methods.

The rank correlation method gave very similar ranking as EVPPI methods. Cost of MI THR was again identified as the variable with the highest importance measure. The top 7 variables based on EVPPI methods were selected to be top 7 based on rank correlation method. The only difference was that Cost of STD THR ranked second and Incremental Utility on 6 months ranked third based on EVPPIs , and vice versa based on rank correlation. However, the rank correlation scores for these 2 variables were very close (raw scores of 0.448 and 0.446 for Cost of MI THR and UI6mnth respectively, or calibrated scores of 79.1 and 78.7 respectively. With the exception of the first three parameters, for the rest of the parameters rank correlation and EVPPI methods did not result with similar calibrated scores.

In terms of computation complexity, rank correlation is very efficient method requiring only a simple data manipulation and one MCS of 5000 replications for all parameters.

However, the interpretation of rank correlation coefficients can be difficult. Coyle argues that one could have a high degree of correlation between a parameter and the outcome but there may be minimal risk of making a wrong decision (Coyle et al 2003). Similarly Claxton noted that simple correlation can be high even when the contribution to the INB variance is low.

The rank correlation method focuses on the degree of similarity between ranking the input and output variables, i.e. the degree that the input variable is responsible for the dispersion in the outcome variable.

Figure 10: Results – Rank Correlation Coefficients

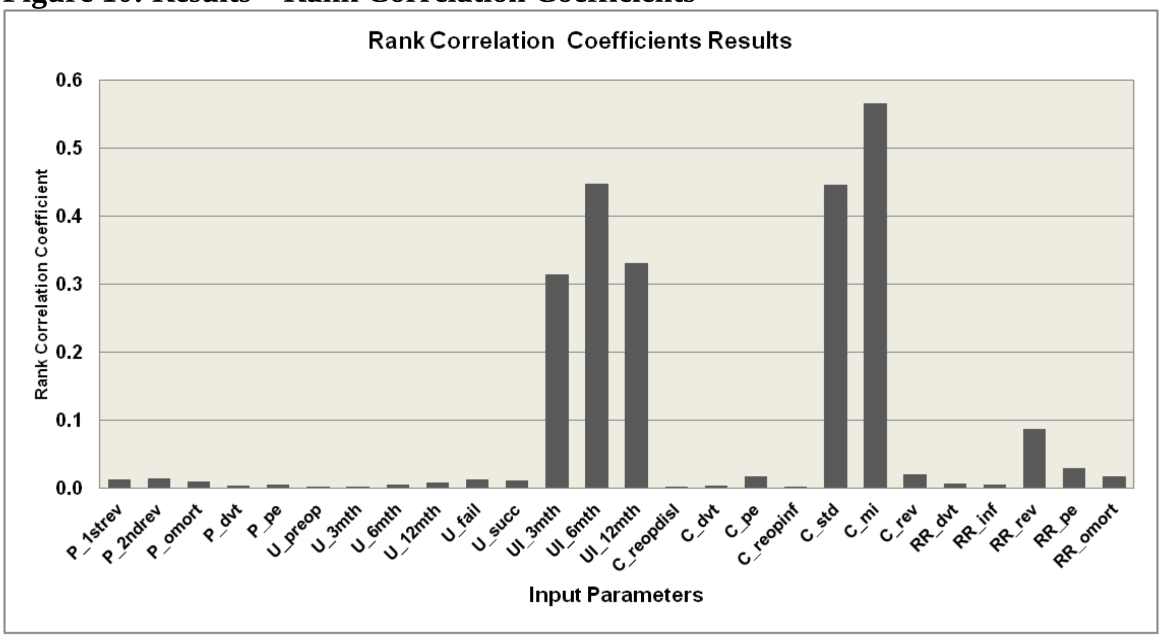


Table 13: Comparison of the Results produced by Rank Correlation Method and EVPPI-Double MCS Method

Parameter	Rank Correlation Coefficient rank (calibrated score)	EVPPI double MCS rank (calibrated score)
U_preop	1 (100)	1 (100)
UI_3mth	2 (79.1)	3 (71.7)

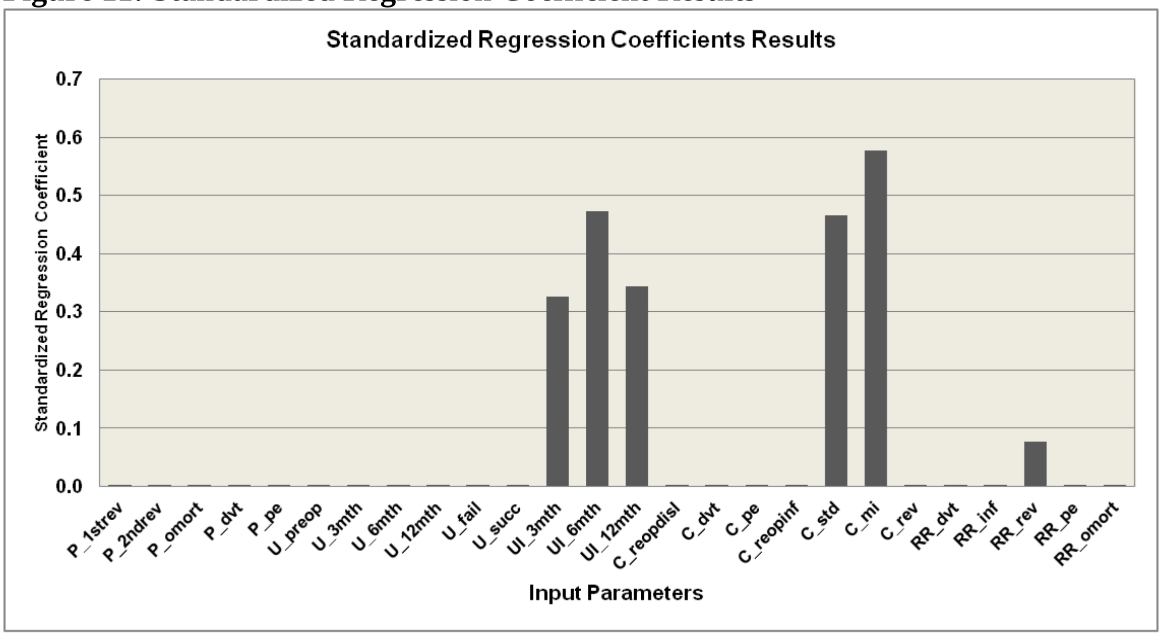
Parameter	Rank Correlation Coefficient rank (calibrated score)	EVPPI double MCS rank (calibrated score)
UI_6mth	3 (78.7)	2 (73.8)
UI_12mth	4 (58.3)	4 (21.5)
U_fail	5 (55.6)	5 (13.8)
U_succ	6 (15.3)	6 (0.3)
P_1strev	7 (5.1)	7 (0)
P_2ndrev	8 (3.4)	7 (0)
P_omort	9 (3.1)	7 (0)
P_dvt	10 (2.9)	7 (0)
P_pe	11 (2.5)	7 (0)
U_3mth	12 (2.2)	7 (0)
U_6mth	13 (2.1)	7 (0)
U_12mth	14 (1.9)	7 (0)
C_reopdisl	15 (1.5)	7 (0)
C_dvt	16 (1.5)	7 (0)
C_pe	17 (1.2)	7 (0)
C_reopinf	18 (0.9)	7 (0)
C_std	19 (0.8)	7 (0)
C_mi	20 (0.7)	7 (0)
C_rev	21 (0.5)	7 (0)
RR_dvt	22 (0.5)	7 (0)
RR_inf	23 (0.1)	7 (0)
RR_rev	24 (0.1)	7 (0)
RR_pe	25 (0.1)	7 (0)
RR_omort	26 (0)	7 (0)

C_dvt= cost of DVT; *C_mi*= cost of MI THR; *C_pe*= cost of PE; *C_reopdisl*=cost of reoperation due to dislocation; *C_reopinf*=cost of reoperation due to infection; *C_rev*= cost of revision THR; *C_std*= cost of STD THR; *EVPPI*= expected value of partial perfect information; *MCS* = Monte Carlo simulation; *P_1strev*=probability of 1st revision; *P_2ndrev*=probability of second revision; *P_dvt*=probability of DVT; *P_omort*=probability of operative mortality; *P_pe*=probability of PE; *RR_dvt*= relative risk of DVT; *RR_inf*= relative risk of infection; *RR_omort*=relative risk of operative mortaility; *RR_pe*= relative risk of PE; *RR_rev*= relative risk of revision; U_3mth= utility at 3 months; *U_6mth*= utility at 6 months; *U_12mth*=utility at 12 months; ; UI_3mth=incremental utility at 3 months; *UI_6mth*= incremental utility at 6 months; *UI_12mth*=incremental utility at 12 months; *U_fail* = utility at failure THR; *U_preop*= preoperative utlity*; U_succ*=utility at success THR

4.3 Standardized Regression Coefficient Method

Results of the standardized regression coefficient method are presented in Figure 11 below, and its calibrated scores are compared with EVPPI in table 14.

Figure 11: Standardized Regression Coefficient Results



Similar to the correlation method, the ranking of the top 6 parameters was consistent with the ranking produced with the EVPPIs method, such that the ranking matched for the 1st, 4th, 5th and 6th parameter, while 2nd and 3rd parameter had reversed rankings. Again, the cost MI THR was identified as a parameter with the highest importance score.

Table 14: Comparison among Calibrated Scores and Ranks of Standardized Regression Coefficient and EVPPI

Parameter	Standardized Correlation Coefficient rank (calibrated score)	EVPPI double MCS rank (calibrated score)
C_mi	1 (100)	1 (100)
UI_6mth	2 (82.1)	3 (71.7)
C_std	3 (80.7)	2 (73.8)
UI_12mth	4 (59.6)	4 (21.5)
UI_3mth	5 (56.6)	5 (13.8)
RR_rev	6 (13.2)	6 (0.3)
RR_omort	7 (0.1)	7 (0)
U_succ	8 (0.1)	7 (0)
U_12mth	9 (0.1)	7 (0)
C_reopdisl	10 (0)	7 (0)
RR_pe	11 (0)	7 (0)
U_fail	12 (0)	7 (0)
P_2ndrev	13 (0)	7 (0)
C_rev	14 (0)	7 (0)
P_dvt	15 (0)	7 (0)
P_pe	16 (0)	7 (0)
U_3mth	17 (0)	7 (0)
C_pe	18 (0)	7 (0)
U_preop	19 (0)	7 (0)
P_1strev	20 (0)	7 (0)
C_reopinf	21 (0)	7 (0)
RR_inf	22 (0)	7 (0)
P_opmort	23 (0)	7 (0)
RRdvt	24 (0)	7 (0)
C_dvt	25 (0)	7 (0)
U_6mth	26 (0)	7 (0)

C_dvt= cost of DVT; *C_mi*= cost of MI THR; *C_pe*= cost of PE; *C_reopdisl*=cost of reoperation due to dislocation; *C_reopinf*=cost of reoperation due to infection; *C_rev*= cost of revision THR; *C_std*= cost of STD THR; *EVPPI*= expected value of partial perfect information; *MCS* = Monte Carlo simulation; *P_1strev*=probability of 1st revision; *P_2ndrev*=probability of second revision; *P_dvt*=probability of DVT; *P_omort*=probability of operative mortality; *P_pe*=probability of PE; *RR_dvt*= relative risk of DVT; *RR_inf*= relative risk of infection; *RR_omort*=relative risk of operative mortaility; *RR_pe*= relative risk of PE; *RR_rev*= relative risk of revision; *SRC* = standardized regression coefficicent; U_3mth= utility at 3 months; *U_6mth*= utility at 6 months; *U_12mth*=utility at 12 months; ; UI_3mth=incremental utility at 3 months; *UI_6mth*= incremental utility at 6 months;

Parameter	Standardized Correlation Coefficient rank (calibrated score)	EVPPI double MCS rank (calibrated score)
UI_12mth=incremental utility at 12 months; *U_fail* = utility at failure THR; *U_preop*= preoperative utlity*; U_succ*=utility at success THR		

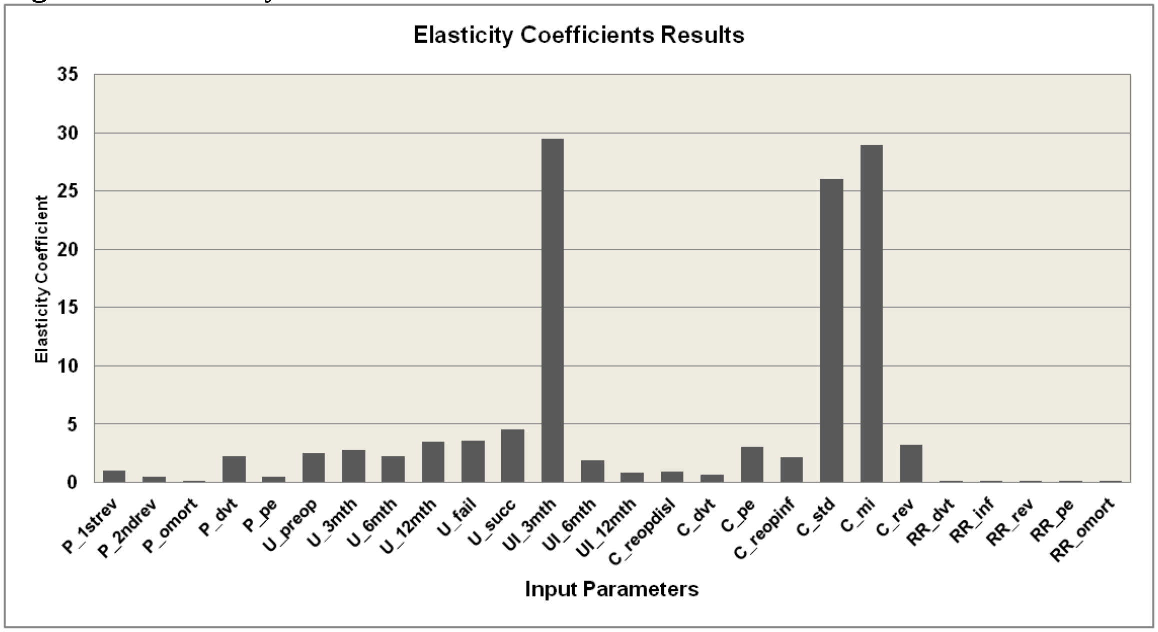
This method was executed in Excel using Excel Add-On: OLS Regression and OLSReg Function. The method is computationally relatively efficient, with a short run time – it requires only one MCS of 5000 replications for all parameters. It requires some data manipulation.

This value of SRC is indicative of the amount of influence the parameter has on the whole model and it does not represent the risk of making a wrong decision.

4.4 Elasticity Coefficient Method

Elasticity coefficients for each parameter are presented in the Figure 12 below.

Figure 12: Elasticity Coefficients Results



The elasticity method gave notably different results than EVPPI methods, as well as correlation and regression coefficients. Incremental utility at 3 months resulted as being the parameter with the highest elasticity coefficient, followed by costs of MI and STD THR. These three parameters had significantly higher elasticity coefficients than the rest of the parameters. Results differ from the ranking based on EVPPIs, with utility of success THR and utility of failure ranked 4^{th} and 5^{th}.

Since the elasticity measure significantly differs from EVPPI, it can be concluded that elasticity coefficient is not an appropriate method for measuring parameters' importance. As expressed before, parameter uncertainty analysis provides one with an insight into which parameters contribute most to variation in model outcomes. By contrast, elasticity analysis shows to what degree model outcomes can be expected to respond to change in input parameter values (Hunter et al 2000).

Elasticity analyses only identify what the potential effect of a change in parameter levels is on model output range (de Kroon et al 2000).

Therefore, it is clear that interpretation of elasticity and EVPPI analyses differ. EVPPI analysis is important for setting research priorities because it identifies parameters where more information in parameter estimation can lead to the greatest increase in precision of predicted INB. Elasticity analysis ranks the importance of parameters based on their potential to influence INB when changed by a set percentage from the mean.

Table 15: Comparison among Calibrated Scores and Ranks of Elasticity Coefficients and EVPPI Results

Parameter	Elasticity Coefficient rank (calibrated score)	EVPPI double MCS rank (calibrated score)
UI_3mth	1 (100)	5 (13.8)
C_mi	2 (98.2)	1 (100)
C_std	3 (88.3)	2 (73.8)
U_succ	4 (15.2)	7 (0)
U_fail	5 (11.9)	7 (0)
U_12mth	6 (11.7)	7 (0)
C_rev	7 (10.8)	7 (0)
C_pe	8 (10)	7 (0)
U_3mth	9 (9.2)	7 (0)
U_preop	10 (8.2)	7 (0)
U_6mth	11 (7.4)	7 (0)
P_dvt	12 (7.2)	7 (0)
C_reopinf	13 (6.9)	7 (0)
UI_6mth	14 (6.3)	3 (71.7)
P_1strev	15 (3.2)	7 (0)
C_reopdisl	16 (3)	7 (0)
UI_12mth	17 (2.7)	4 (21.5)
C_dvt	18 (2)	7 (0)
P_2ndrev	19 (1.3)	7 (0)
P_pe	20 (1.2)	7 (0)
RR_rev	21 (0.2)	6 (0.3)
RR_opmort	22 (0.2)	7 (0)
RR_inf	23 (0.1)	7 (0)
RR_pe	24 (0.1)	7 (0)
RR_dvt	25 (0.1)	7 (0)
P_opmort	26 (0)	7 (0)

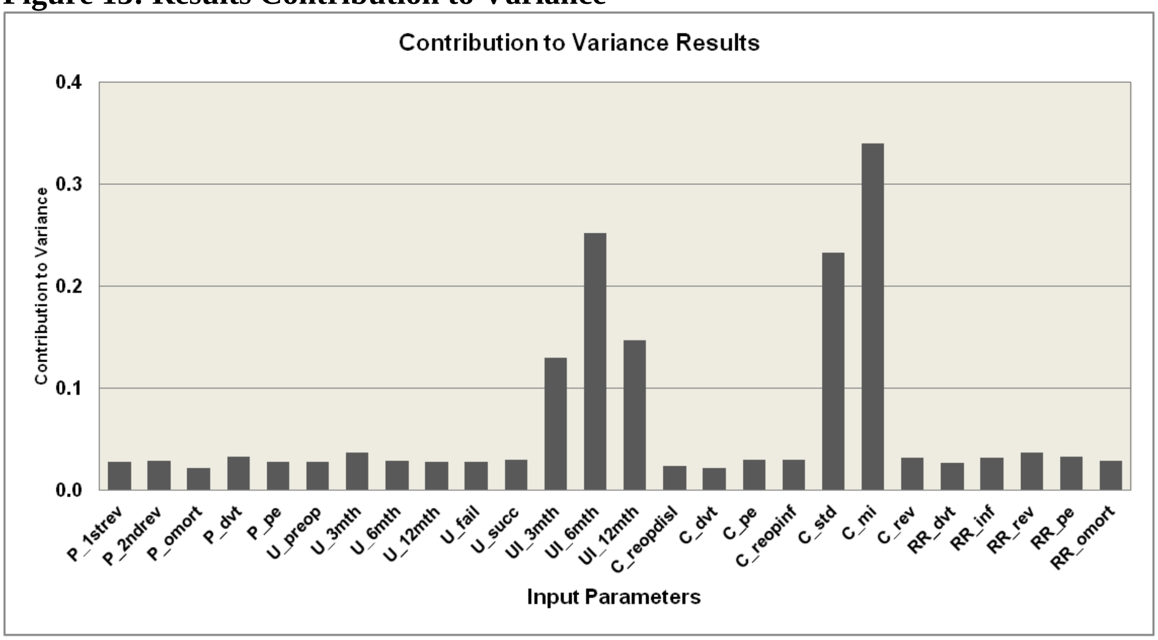
C_dvt= cost of DVT; *C_mi*= cost of MI THR; *C_pe*= cost of PE; *C_reopdisl*=cost of reoperation due to dislocation; *C_reopinf*=cost of reoperation due to infection; *C_rev*= cost of revision THR; *C_std*= cost of STD THR; *EVPPI*= expected value of partial perfect information; *MCS* = Monte Carlo simulation; *P_1strev*=probability of 1st revision; *P_2ndrev*=probability of second revision; *P_dvt*=probability of DVT; *P_omort*=probability of operative mortality; *P_pe*=probability of PE; *RR_dvt*= relative risk of DVT; *RR_inf*= relative risk of infection; *RR_omort*=relative risk of operative mortaility; *RR_pe*= relative risk of PE; *RR_rev*= relative risk of revision; U_3mth= utility at 3 months; *U_6mth*= utility at 6 months; *U_12mth*=utility at 12 months; ; UI_3mth=incremental utility at 3 months; *UI_6mth*= incremental utility at 6 months; *UI_12mth*=incremental utility at 12 months; *U_fail* = utility at failure THR; *U_preop*= preoperative utlity; *U_succ*=utility at success THR

In terms of computational complexity, elasticity method was shown to be computationally complex, as it required 505,000 replications for each parameter.

4.5 Contribution to Variance Method

Contribution to variance measure for each input parameter was calculated and results are presented in Figure 13 below.

Figure 13: Results Contribution to Variance



The contributions of variance measure highly conform with regression and correlation measures. Specifically, the top five parameters are again selected to be cost of STD and MI THR parameters, as well as the incremental utilities for 3 ,6 and 12 months. The rest of the parameters have notably lower contribution to variance, which can be observed in the table below showing the calibrated scores as well as the ranking.

This method has equivalent computational complexity as the elasticity measure, where by using Simpsons' rule for approximation of integration, one can estimate contribution to variance measures with 505,000 replications per parameter.

Table 16 compares the ranks and calibrated scores of contribution to variance coefficient and EVPPI.

Table 16: Comparison among Calibrated Scores and Ranks of Contribution to Variance and EVPPI Results

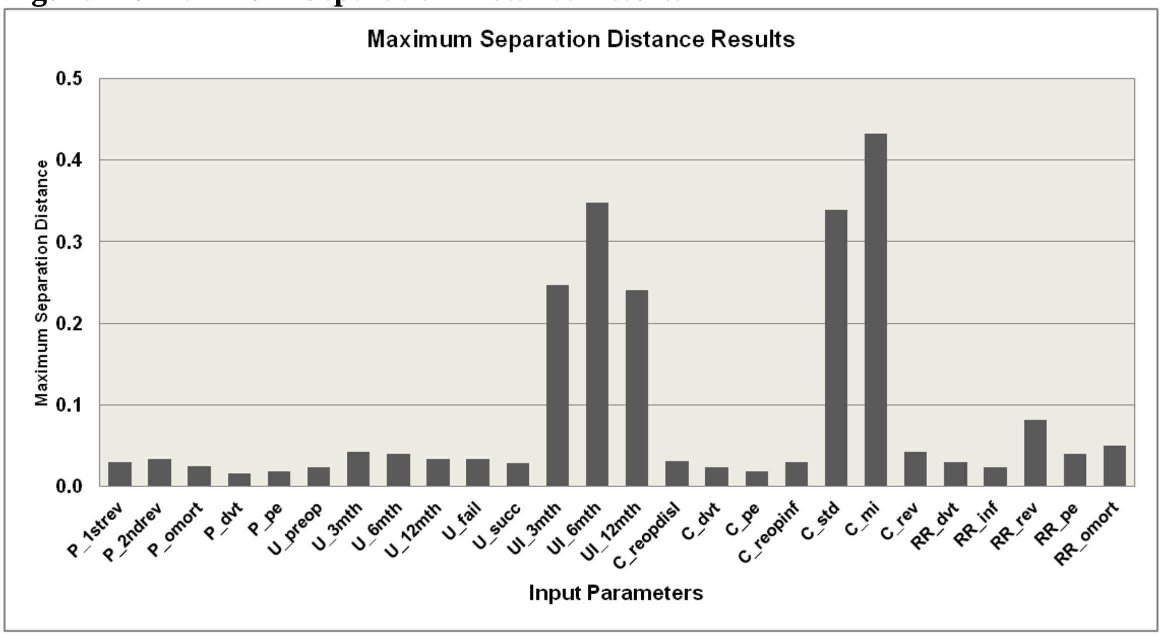
Parameter	Contribution to variance rank (calibrated score)	EVPPI double MCS rank (correlation score)
C_mi	1 (100)	1 (100)
UI_6mth	2 (72.5)	3 (71.7)
C_std	3 (66.5)	2 (73.8)
UI_12mth	4 (39.5)	4 (21.5)
UI_3mth	5 (33.8)	5 (13.8)
RR_rev	6 (4.6)	6 (0.3)
U3_mth	7 (4.6)	7 (0)
P_dvt	8 (3.4)	7 (0)
RR_pe	9 (3.3)	7 (0)
RR_inf	10 (3.1)	7 (0)
C_rev	11 (3)	7 (0)
C_reopinf	12 (2.3)	7 (0)
U_succ	13 (2.3)	7 (0)
C_pe	14 (2.3)	7 (0)
U_6mth	15 (2.3)	7 (0)
P_2ndrev	16 (2.1)	7 (0)
RR_omort	17 (2.1)	7 (0)
U_12mth	18 (2)	7 (0)
U_fail	19 (1.9)	7 (0)
P_pe	20 (1.9)	7 (0)
U_preop	21 (1.8)	7 (0)
P_1strev	22 (1.7)	7 (0)
RR_dvt	23 (1.4)	7 (0)
C_reopdisl	24 (0.6)	7 (0)
C_dvt	25 (0)	7 (0)
P_omort	26 (0)	7 (0)

C_dvt= cost of DVT; *C_mi*= cost of MI THR; *C_pe*= cost of PE; *C_reopdisl*=cost of reoperation due to dislocation; *C_reopinf*=cost of reoperation due to infection; *C_rev*= cost of revision THR; *C_std*= cost of STD THR; *EVPPI*= expected value of partial perfect information; *MCS* = Monte Carlo simulation; *P_1strev*=probability of 1st revision; *P_2ndrev*=probability of second revision; *P_dvt*=probability of DVT; *P_omort*=probability of operative mortality; *P_pe*=probability of PE; *RR_dvt*= relative risk of DVT; *RR_inf*= relative risk of infection; *RR_omort*=relative risk of operative mortaility; *RR_pe*= relative risk of PE; *RR_rev*= relative risk of revision; U_3mth= utility at 3 months; *U_6mth*= utility at 6 months; *U_12mth*=utility at 12 months; ; UI_3mth=incremental utility at 3 months; *UI_6mth*= incremental utility at 6 months; *UI_12mth*=incremental utility at 12 months; *U_fail* = utility at failure THR; *U_preop*= preoperative utlity; *U_succ*=utility at success THR

4.6 Maximum Separation Distance Method

The results of maximum separation distance method are presented in the Figure 14 below.

Figure 14: Maximum Separation Distance Results

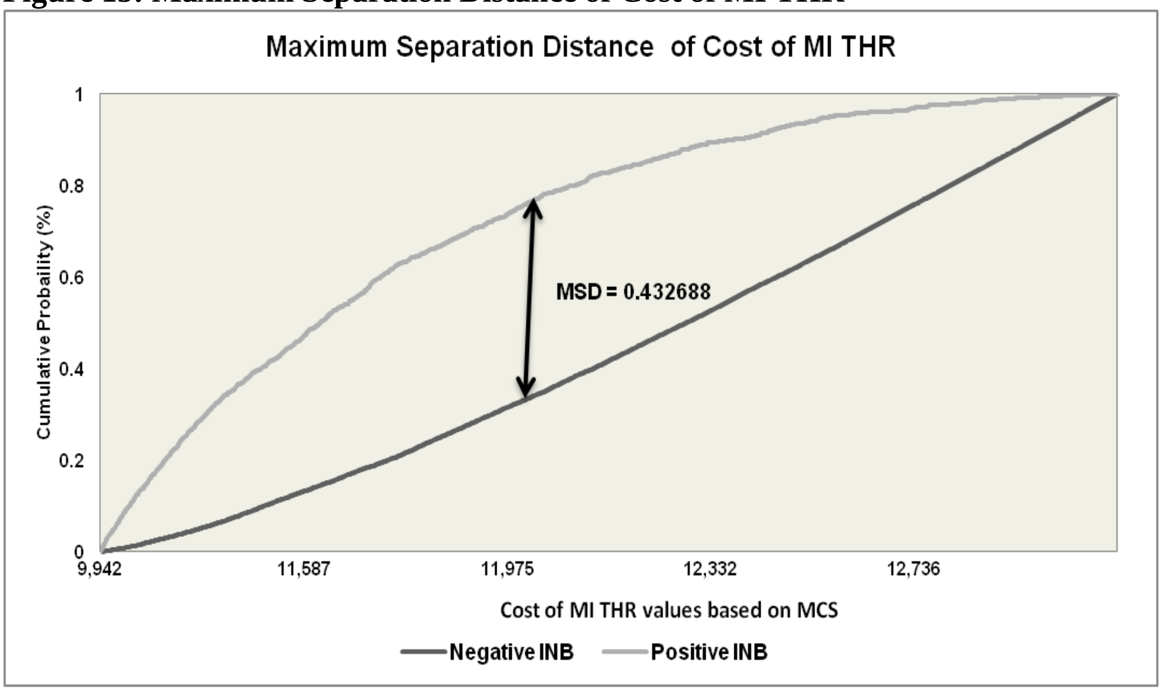


The MSD results are largely conforming to the EVPPI results too, as shown in Table 17. All six parameters that have non-zero EVPPIs were selected to be top 6 parameters based on MSD ranking. Cost of MI THR remains to be the most important parameter. Similarly to all the other alternative measures previously assessed, incremental utility at 6 months emerged as 2nd important parameter, followed by cost of STD THR. Incremental utilities at 3 months and 12 months had similar calibrated scales and ranked 4th and 5th. It can also be noted that MSD values for parameters that have zero EVPPIs are substantially lower than the MSDs for parameters with non-zero EVPPI.

As Coyle remarked, this is likely because maximum separation distances focus on the uncertainty concerning the optimal decision based on the outcome measure.

An example of the concept of MSD is presented in the Figure 15.

Figure 15: Maximum Separation Distance of Cost of MI THR



Computationally, MSD is efficient method, requiring one MCS of 5,000 replications for all variables and relatively simple calculation that can be executed in MS Excel, i.e. no additional software was required.

Table 17: Comparison among Calibrated Scores and Ranks of Maximum Separation Distance and EVPPI Results

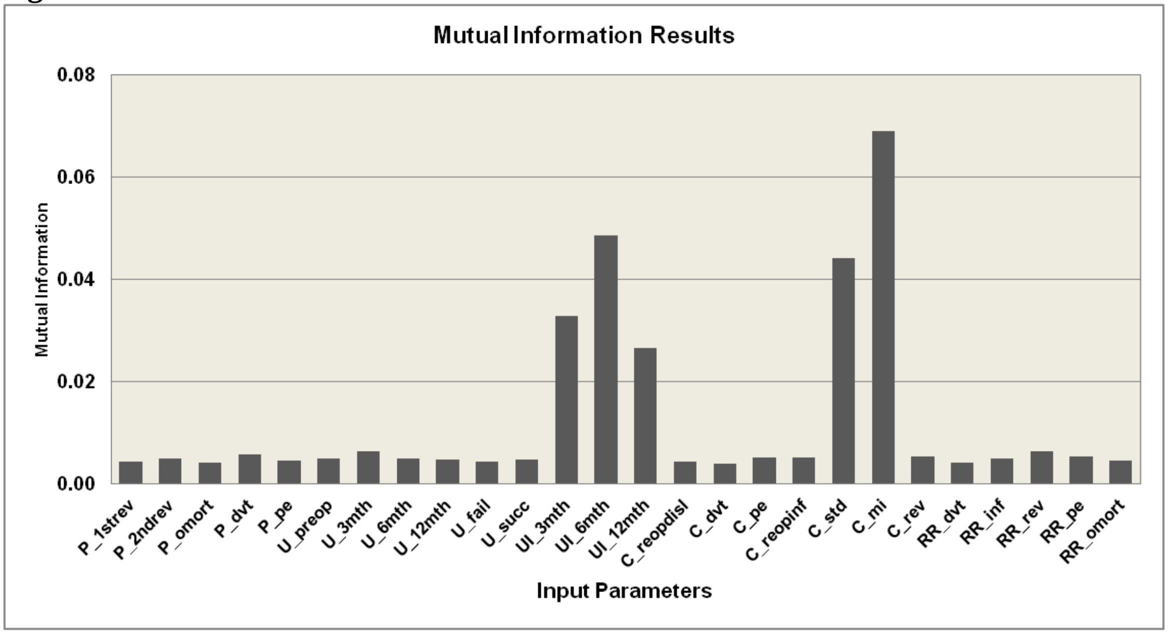
Parameters	Maximum Separation Distance rank (calibrated score)	EVPPI double MCS rank (calibrated score)
C_mi	1 (100)	1 (100)
UI_6mth	2 (79.7)	3 (71.7)
C_std	3 (77.6)	2 (73.8)
UI_3mth	4 (55.6)	5 (13.8)
UI_12mth	5 (53.9)	4 (21.5)
RR_rev	6 (15.8)	6 (0.3)
RR_omort	7 (8.3)	7 (0)
U_3mth	8 (6.4)	7 (0)
C_rev	9 (6.4)	7 (0)
RR_pe	10 (5.9)	7 (0)
U_6mth	11 (5.8)	7 (0)
U_12mth	12 (4.5)	7 (0)
P_2ndrev	13 (4.4)	7 (0)
U_fail	14 (4.3)	7 (0)
C_reopdisl	15 (3.9)	7 (0)
RR_dvt	16 (3.5)	7 (0)
C_reopinf	17 (3.5)	7 (0)
P_1strev	18 (3.4)	7 (0)
U_succ	19 (3.3)	7 (0)
P_omort	20 (2.4)	7 (0)
RR_inf	21 (2)	7 (0)
U_preop	22 (1.9)	7 (0)
C_dvt	23 (1.8)	7 (0)
C_pe	24 (0.8)	7 (0)

Parameters	Maximum Separation Distance rank (calibrated score)	EVPPI double MCS rank (calibrated score)
P_pe	25 (0.7)	7 (0)
P_dvt	26 (0)	7 (0)
C_dvt= cost of DVT; *C_mi*= cost of MI THR; *C_pe*= cost of PE; *C_reopdisl*=cost of reoperation due to dislocation; *C_reopinf*=cost of reoperation due to infection; *C_rev*= cost of revision THR; *C_std*= cost of STD THR; *EVPPI*= expected value of partial perfect information; *MCS* = Monte Carlo simulation; *P_1strev*=probability of 1st revision; *P_2ndrev*=probability of second revision; *P_dvt*=probability of DVT; *P_omort*=probability of operative mortality; *P_pe*=probability of PE; *RR_dvt*= relative risk of DVT; *RR_inf*= relative risk of infection; *RR_omort*=relative risk of operative mortaility; *RR_pe*= relative risk of PE; *RR_rev*= relative risk of revision; U_3mth= utility at 3 months; *U_6mth*= utility at 6 months; *U_12mth*=utility at 12 months; ; UI_3mth=incremental utility at 3 months; *UI_6mth*= incremental utility at 6 months; *UI_12mth*=incremental utility at 12 months; *U_fail* = utility at failure THR; *U_preop*= preoperative utlity*; U_succ*=utility at success THR		

4.7. Mutual Information

The results of the mutual information method are presented in Figure 16.

Figure 16: Mutual Information Results



Results of MI measure are consistent with the other alternative measures, as well as with EVPPI results. As observed, cost of STD and MI THR as well as incremental utilities have the highest MI value.

This method has equivalent computational complexity as elasticity and contribution to variance measure, where using Simpsons' rule for approximation of integration, one can estimate contribution to variance measures with 505,000 replications per parameter.

Table 18: Comparison among Calibrated Scores and Ranks of Mutual Information and EVPPI Results

Parameters	Mutual Information rank (calibrated score)	EVPPI double-MCS Rank (calibrated score)
C_mi	1 (100)	1 (100)
UI_6mth	2 (68.7)	3 (71.7)

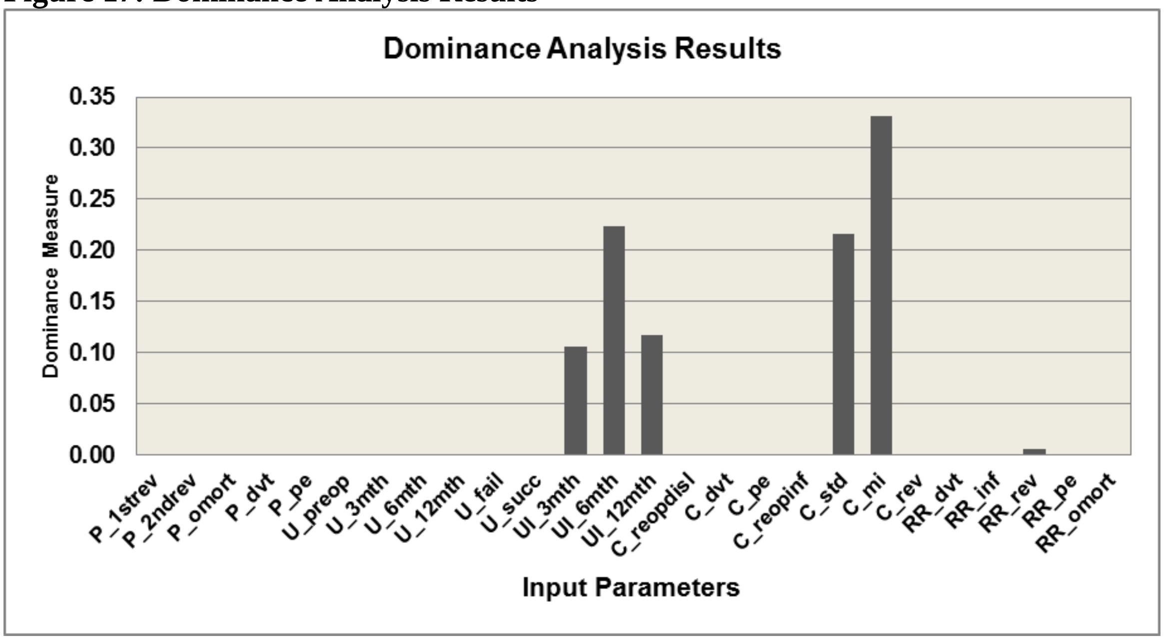
Parameters	Mutual Information rank (calibrated score)	EVPPI double-MCS Rank (calibrated score)
C_std	3 (61.8)	2 (73.8)
UI_3mth	4 (44.5)	5 (13.8)
UI_12mth	5 (34.6)	4 (21.5)
RR_rev	6 (3.7)	6 (0.3)
U_3mth	7 (3.6)	7 (0)
P_dvt	8 (2.7)	7 (0)
C_rev	9 (2.2)	7 (0)
RR_pe	10 (1.9)	7 (0)
C_pe	11 (1.7)	7 (0)
C_reopinf	12 (1.7)	7 (0)
RR_inf	13 (1.6)	7 (0)
U_6mth	14 (1.4)	7 (0)
U_preop	15 (1.4)	7 (0)
P_2ndrev	16 (1.3)	7 (0)
U_12mth	17 (1.2)	7 (0)
U_succ	18 (1)	7 (0)
P_pe	19 (1)	7 (0)
RR_omort	20 (0.7)	7 (0)
U_fail	21 (0.6)	7 (0)
P_1strev	22 (0.6)	7 (0)
C_reopdisl	23 (0.4)	7 (0)
RR_dvt	24 (0.3)	7 (0)
P_omort	25 (0.1)	7 (0)
C_dvt	26 (0)	7 (0)

C_dvt= cost of DVT; *C_mi*= cost of MI THR; *C_pe*= cost of PE; *C_reopdisl*=cost of reoperation due to dislocation; *C_reopinf*=cost of reoperation due to infection; *C_rev*= cost of revision THR; *C_std*= cost of STD THR; *EVPPI*= expected value of partial perfect information; *MCS* = Monte Carlo simulation; *P_1strev*=probability of 1st revision; *P_2ndrev*=probability of second revision; *P_dvt*=probability of DVT; *P_omort*=probability of operative mortality; *P_pe*=probability of PE; *RR_dvt*= relative risk of DVT; *RR_inf*= relative risk of infection; *RR_omort*=relative risk of operative mortaility; *RR_pe*= relative risk of PE; *RR_rev*= relative risk of revision; U_3mth= utility at 3 months; *U_6mth*= utility at 6 months; *U_12mth*=utility at 12 months; ; UI_3mth=incremental utility at 3 months; *UI_6mth*= incremental utility at 6 months; *UI_12mth*=incremental utility at 12 months; *U_fail* = utility at failure THR; *U_preop*= preoperative utlity*;* *U_succ*=utility at success THR

4.8. Dominance Analysis Method

The results of dominance measure are summarized in Figure 17 and a comparison to the EVPPI results is given in Table 19.

Figure 17: Dominance Analysis Results



As opposed to the other alternative measures, dominance measure resulted with non-zero values only for 6 out of 26 parameters, specifically the same 6 parameters selected with EVPPI method. Moreover, the calibrated scores for dominance measure are very similar to the EVPPI values based on the double-MCS method. However, as with MSD and rank correlation, the dominance measure method is computationally highly more efficient than the double-MCS method, since it requires a sample of data obtained from only one MCS of 5,000 replications.

Table 19: Comparison among Calibrated Scores and Ranks of Dominance Measure and EVPPI Results for the Parameters with non-zero Dominance Measure and EVPPI

Parameter^	Dominance Measure rank (calibrated score)	EVPPI Double MCS rank (calibrated score)
C_mi	1 (100)	1 (100)
UI_6mth	2 (67.3)	3 (71.7)
C_std	3 (65.2)	2 (73.8)
UI_12mth	4 (35.5)	4 (21.5)
UI_3mth	5 (32)	5 (13.8)
RR_rev	6 (1.8)	6 (0.3)
C_mi= cost of |MI THR; *C_std*=cost of STD THR; *EVPP*I= expected value of partial perfect information; *MCS* = Monte Carlo simulation *RR_rev* = relative risk of revision; *UI_3mth*= incremental utility at 3 months; *UI_6mth*= incremental utility at 6 months; *UI_12mth*= incremental utility at 12 months		

^Only parameters with non-zero dominance measure and EVPPI are included. The rest of the parameters resulted with zero dominance measure and EVPPI.

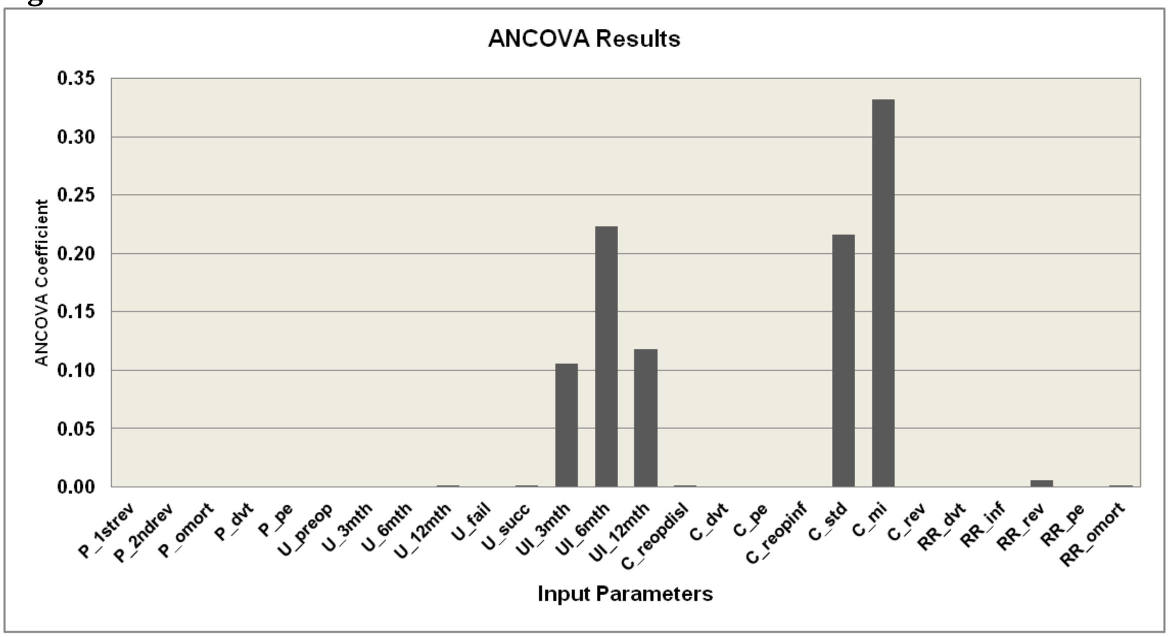
4.9. ANCOVA

Analysis of co-variance i.e. ANCOVA was also performed to identify the most important input parameters in the model. Results from the analysis in terms of ANCOVA coefficients are summarized in Figure 18. Also, as with any other techniques, there is a table below comparing the results of ANCOVA with EVPPI methods. (Table 20)

The standardized ANCOVA coefficients had exactly same ranking as the dominance measure and standardized regression coefficients techniques, i.e. costs of STD and MI

THR, followed by incremental utilities and the risk of revision were selected to be the most important parameters.

Figure 18: ANCOVA Results



Computationally, ANCOVA is an efficient method to be used as it requires only one MCS of 5,000 replications. In this thesis, ANCOVA was conducted using an additional statistical package: the XLSTAT add-on to MS EXCEL.

The ANCOVA analysis also provides additional information on the input parameters as multicolinearity between the input parameters can be assessed. Multicollinearity is a measure of the relationship between the regression exploratory variables. Whenever two supposedly independent variables are highly correlated, it will be difficult to assess their relative importance in determining some dependent variable. The higher the correlation between independent variables, the greater the sampling error of the partials (Blalock 1963).

Table 20: Comparison among Calibrated Scores and Ranks of ANCOVA and EVPPI Results for the Parameters with non-zero ANCOVA and EVPPI

Parameters^	ANCOVA coefficient rank (calibrated score)	EVPP double MCS rank (calibrated score)
C_mi	1 (100)	1 (100)
UI_6mth	2 (82.1)	3 (71.7)
C_std	3 (80.7)	2 (73.8)
UI_12mth	4 (59.6)	4 (21.5)
UI_3mth	5 (56.6)	5 (13.8)
RR_rev	6 (13.2)	6 (0.3)
RR_omort	7 (0.1)	7 (0)
U_succ	8 (0.1)	7 (0)
U_12mth	9 (0.1)	7 (0)

C_mi= cost of |MI THR; *C_std*=cost of STD THR; *EVPPI*= expected value of partial perfect information; *MCS* = Monte Carlo simulation; *RR_omort*=relative risk of operative mortality; *RR_rev* = relative risk of revision; *UI_3mth*= incremental utility at 3 months; *UI_6mth*= incremental utility at 6 months; *UI_12mth*= incremental utility at 12 months; *U_succ* = utility for success THR

^Only parameters with non-zero ANCOVA and EVPPI are included. The rest of the parameters resulted with zero ANCOVA and EVPPI.

Tolerance and the Variance Inflation Factor (VIF) are two collinearity diagnostic factors that can help identify multicollinearity.

$$VIF(\beta_i) = \frac{1}{1 - {R_i}^2} \quad Tolerance(\beta_i) = \frac{1}{VIF(\beta_i)}$$

The higher VIF or the lower the tolerance index, the higher the variance of β_i and the greater the chance of finding β_i insignificant,which means that severe multicolinearity effects are present (Greene 2000). A small tolerance value indicates that the variable under consideration is almost a perfect linear combination of the independent variables already in the equation and that it should not be added to the regression equation. Results of multicolinearity in the hip replacement model are presented in the table 21 below.

Table 21: Tolerance and the Variance Inflation Factor

Parameter	Tolerance	VIF
U_preop	0.994	1.006
U_3mth	0.998	1.002
U_6mth	0.996	1.004
U_12mth	0.993	1.008
P_succ	0.997	1.003
P_rev	0.995	1.005
P_1strev	0.997	1.003
P_2ndrev	0.996	1.004
P_omort	0.996	1.004
P_dvt	0.998	1.002
P_pe	0.996	1.004
UI_3mth	0.993	1.007
UI_6mth	0.996	1.004
UI_12mth	0.997	1.003
C_reopdisl	0.994	1.006
C_dvt	0.996	1.004
C_pe	0.994	1.006
C_reopdisl	0.996	1.004
C_std	0.996	1.004
C_mi	0.997	1.003
C_rev	0.996	1.004
RR_dvt	0.995	1.005
RR_inf	0.995	1.005
RR_inf	0.994	1.006
RR_pe	0.997	1.003
RR_omort	0.993	1.007

C_dvt= cost of DVT; *C_mi*= cost of MI THR; *C_pe*= cost of PE; *C_reopdisl*=cost of reoperation due to dislocation; *C_reopinf*=cost of reoperation due to infection; *C_rev*= cost of revision THR; *C_std*=

Parameter	Tolerance	VIF
cost of STD THR; *P_1strev*=probability of 1st revision; *P_2ndrev*=probability of second revision; *P_dvt*=probability of DVT; *P_omort*=probability of operative mortality; *P_pe*=probability of PE; *RR_dvt*= relative risk of DVT; *RR_inf*= relative risk of infection; *RR_omort*=relative risk of operative mortaility; *RR_pe*= relative risk of PE; *RR_rev*= relative risk of revision; U_3mth= utility at 3 months; *U_6mth*= utility at 6 months; *U_12mth*=utility at 12 months; ; UI_3mth=incremental utility at 3 months; *UI_6mth*= incremental utility at 6 months; *UI_12mth*=incremental utility at 12 months; *U_fail* = utility at failure THR; *U_preop*= preoperative utlity; *U_succ*=utility at success THR; VIF = variance inflation factor		

The results clearly show that multicolinearity is not an issue in this model.

4.10. Sequential Bifurcation

Figure 19 represents the step-by step approach in conducting the sequential bifurcation analysis, revealing that in 21 steps, 5 out of 26 parameters were selected as important parameters.

Figure 19: Sequential Bifurcation, displays the selection of important parameters in step-by-step approach

Run	Var 1	Var 2	Var 3	Var 4	Var 5	Var 6	Var 7	Var 8	Var 9	Var 10	Var 11	Var 12	Var 13	Var 14	Var 15	Var 16	Var 17	Var 18	Var 19	Var 20	Var 21	Var 22	Var 23	Var 24	Var 25	Var 26
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
6	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

1	**Switched-on parameters**
1	**Important parameter**
0	**Non-imporant parameter**

The β parameter values obtained for each step of the SB process are presented in the Table 22 below. The threshold B was not explicitly introduced, however the large difference among β values made it easy to conclude which parameters should be selected as important.

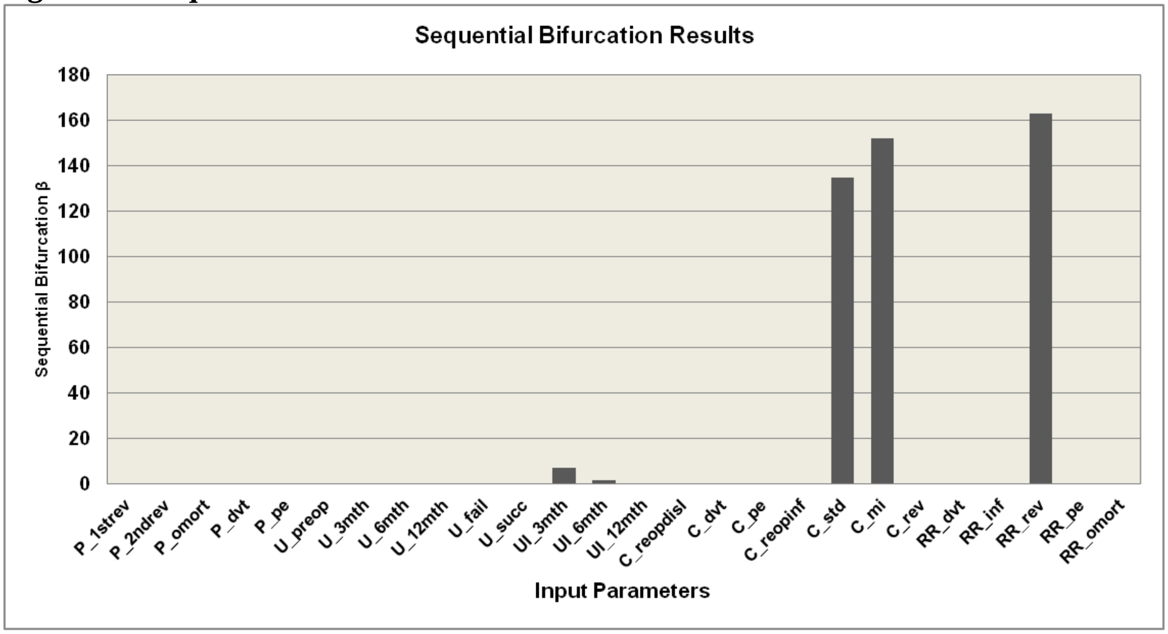
Table 22: β Values for Each Sequential Bifurcation Step

SB step	Switched on parameters	***β***
step 1	None	0.00
step 2	All	451.61
step 3	Var 1- Var 11	0.23
step 4	Var 12-Var 18	286.58
step 5	Var 19- Var 26	172.58
step 6	Var 12-Var	0.00

SB step	Switched on parameters	β
	15	
step 7	Var 16-Var 18	286.58
step 8	None	0.00
step 9	Var 16-Var 17	286.58
step 10	Var 16	134.57
step 11	Var 17	152.01
step 12	Var 18	0.00
step 13	Var19-Var22	8.79
step 14	Var 21-Var22	0.04
step 15	Var19	7.03
step 16	Var 20	1.72
step 17	Var 23-Var 24	162.80
step 18	Var 23	0.01
step 19	Var 24	162.79
step 20	Var 25-Var 26	0.99
step 21	None	0

The results of sequential bifurcation method are presented in the Figure 20.

Figure 20: Sequential Bifurcation Results



As opposed to EVPPI methods, SB resulted in the highest importance attributed to relative risk of revision, opposed to cost of MI and STD THR based on the EVPPI techniques. Cost of STD and MI THR scored 2nd and 3rd. Incremental utilities at 3 months and 6 months were also identified as parameters with non-zero importance, but significantly lower than the relative risk of revision, cost of STD and cost of MI THR.

It should be noted that a priori knowledge of input parameters based on the previous methods was used to define subgroups of interest, such that the group of parameters that are expected not to be important could be eliminated in the earlier stages.

Table 23: Comparison among Calibrated Scores and Ranks of Sequential Bifurcation and EVPPI Results for the Parameters with non-zero Sequential Bifurcation Coefficient and EVPPI

Parameter^	Sequential Bifurcation Coefficient rank (calibrated score)	EVPPI double MCS rank (calibrated score)
RR_rev	1 (100)	6 (0.3)
C_mi	2 (93.3)	1 (100)
C_std	3 (82.4)	2 (73.8)
UI_3mth	4 (3.3)	5 (13.8)
UI_6mth	5 (0.1)	3 (71.7)
UI_12mth	6 (0)	4 (21.5)
C_mi= cost of |MI THR; *C_std*=cost of STD THR; *EVPP*I= expected value of partial perfect information; *MCS* = Monte Carlo simulation *RR_rev* = relative risk of revision; *UI_3mth*= incremental utility at 3 months; *UI_6mth*= incremental utility at 6 months; *UI_12mth*= incremental utility at 12 months		

^ Only parameters with non-zero sequential bifurcation coefficient and EVPPI are included. The rest of the parameters resulted with zero sequential bifurcation coefficient and EVPPI.

If the analyst has no prior knowledge of the potential importance of the parameters of interest, then the analysis is likely to require significantly larger number of repetitions.

Also, for models with relatively small number of input parameters, the advantage of creating subgroups of several parameters as opposed to testing individual parameters is uncertain.

4.11. Comparison of Results

Twelve different methods of assessing importance of input parameters were applied to the input parameters of the health economic model comparing the cost- utility of minimally invasive hip replacement in the management of hip arthritic disease to the standard hip replacement. Three out of twelve methods were methods for calculation of EVPPI and 9 were alternative methods.

Two tables below (Table 24 and Table 25) summarize the ranking of parameters by all twelve methods, as well as the calibrated scores on a scale of 0 to 100.

The three EVPPI methods gave similar results, by identifying the same first five parameters as the most important ones (cost of MI THR and STD THR, the incremental utilities at 3, 6 and 12 months). EVPPI double MCS method and quadrature method both identified relative risk of revision as the sixed most important parameter, as opposed to the single MCS method which did not identify the relative risk of revision as an important parameter. Contribution to variance, rank correlation, dominance measure, ANCOVA and standardized regression coefficients had very similar ranking as EVPPI methods. Based on these methods, cost of MI THR and STD THR, the incremental utilities at 3, 6 and 12 months and the relative risk of revision were identified as the six most important measures, only with the reverse order of the second and third parameter. The other two methods (elasticity measure and sequential bifurcation) gave different rankings for the first six parameters.

When comparing the results in terms of calibrated scores, EVPPI double MCS method resulted with higher calibrated scores than EVPPI quadrature and single MCS method. The three EVPPI methods resulted with scores of exactly zero for all parameters except the first six parameters. Among the alternative measures, this feature was observed only with the dominance measure, although ANCOVA and standardized regression coefficients were quite low for the non-first six parameters. As another observation, based on EVPPI double-MCS method, the second and third important parameters had significantly higher calibrated scores than the forth parameter. A difference among third and fourth parameters was as pronounced as in EVPPI double MCS method only with the dominance measure.

Table 24: Ranking of the Parameters Based on Twelve Methods of Measures of Parameters' Importance

Parameter	EVPPI double MCS	EVPPI single MCS	EVPPI -quad	CV	MI	EC	RC	DM	ANCOVA	SRC	MSD	SB
C_mi	1	1	1	1	1	2	1	1	1	1	1	2
C_std	2	2	2	3	3	3	3	3	3	3	3	3
UI_6mth	3	3	3	2	2	14	2	2	2	2	2	5
UI_12mth	4	4	4	4	5	17	4	4	4	4	5	6
UI_3mth	5	5	5	5	4	1	5	5	5	5	4	4
RR_rev	6	6	6	6	6	21	6	6	6	6	6	1
U_preop	7	6	7	21	15	10	24	19	19	19	20	6
U_3mth	7	6	7	7	7	9	23	17	17	17	7	6
U_6mth	7	6	7	15	14	11	19	26	26	26	9	6
U_12mth	7	6	7	18	17	6	16	9	9	10	10	6
U_fail	7	6	7	19	21	5	12	12	12	12	12	6
U_succ	7	6	7	13	18	4	14	8	8	9	17	6

Parameter	EVPPI double MCS	EVPPI single MCS	EVPPI -quad	CV	MI	EC	RC	DM	ANCOVA	SRC	MSD	SB
P_1strev	7	6	7	22	22	15	13	20	20	20	16	6
P_2ndrev	7	6	7	16	16	19	11	13	13	13	11	6
P_omort	7	6	7	26	25	26	15	23	23	23	18	6
P_dvt	7	6	7	8	8	12	22	15	15	15	24	6
P_pe	7	6	7	20	19	20	20	16	16	16	23	6
C_reopdisl	7	6	7	24	23	16	26	10	10	11	13	6
C_dvt	7	6	7	25	26	18	21	25	25	25	21	6
C_pe	7	6	7	14	11	8	9	18	18	18	22	6
C_reopinf	7	6	7	12	12	13	25	21	21	21	15	6
C_rev	7	6	7	11	9	7	8	14	14	14	8	6
RR_dvt	7	6	7	23	24	25	17	24	24	24	14	6
RR_inf	7	6	7	10	13	23	18	22	22	22	19	6
RR_P	7	7	7	10	24	7	11	11	11	25	6	6

Parameter	EVPPI double MCS	EVPPI single MCS	EVPPI -quad	CV	MI	EC	RC	DM	ANCOVA	SRC	MSD	SB
E RR_omort	7	7	7	20	22	10	7	7	7	25	6	6

CV=Contribution to variance, DM= dominance measure; EC=elasticity coefficient ,EVPPI= expected value of partial perfect information; MCS= Monte Carlo simulation; MI=mutual information, MSD=maximum separation distance, RC=rank correlation, SB=sequential bifurcation, SRC=standardized regression coefficient

Table 25: Calibrated Scores of the Parameters Based on Twelve Methods of Measures of Parameters' Importance

Parameters	EVPPI double MCS	EVPPI single MCS	EVPPI quad	CV	MI	EC	RC	DM	ANCOVA	SRC	MSD	SB
C_mi	100	100	100	100	100	98.19	100	100	100	100	100	93.31
C_std	73.76	58.11	55.76	66.5	61.78	88.31	78.73	65.19	80.75	80.75	78.43	82.48
UI_6mth	71.75	47.4	13.58	72.54	68.69	6.25	79.1	67.34	82.07	82.07	80.4	0
UI_12mth	21.48	7.67	10.33	39.49	34.65	2.68	58.3	35.53	59.61	59.61	55.52	N/A
UI_3mth	13.83	4.11	4.84	33.85	44.46	100	55.57	31.96	56.63	56.63	57.14	3.3

Parameters	EVPPI double MCS	EVPPI single MCS	EVPPI quad	CV	MI	EC	RC	DM	ANCOVA	SRC	MSD	SB
RR_rev	0.29	0	0.3	4.65	3.69	0.21	15.26	1.75	13.25	13.25	18.78	100
U_preop	0	0	0	1.84	1.38	8.21	0.11	0	0.02	0.02	5.32	N/A
U_3mth	0	0	0	4.6	3.65	9.16	0.12	0	0.02	0.02	9.7	N/A
U_6mth	0	0	0	2.26	1.39	7.44	0.79	0	0	0	9.13	N/A
U_12mth	0	0	0	1.96	1.24	11.67	1.45	0	0.05	0.05	7.82	N/A
U_fail	0	0	0	1.88	0.63	11.86	2.25	0	0.03	0.03	7.62	N/A
U_succ	0	0	0	2.32	1.01	15.21	1.93	0	0.06	0.06	6.7	N/A
P_1strev	0	0	0	1.69	0.63	3.21	2.08	0	0.01	0.01	6.82	N/A
P_2ndrev	0	0	0	2.07	1.35	1.34	2.55	0	0.03	0.03	7.74	N/A
P_omort	0	0	0	0	0.12	0	1.5	0	0.01	0.01	5.83	N/A
P_dvt	0	0	0	3.36	2.71	7.25	0.49	0	0.02	0.02	3.51	N/A

Parameters	EVPPI double MCS	EVPPI single MCS	EVPPI quad	CV	MI	EC	RC	DM	ANCOVA	SRC	MSD	SB
P_pe	0	0	0	1.87	0.96	1.21	0.67	0	0.02	0.02	4.22	N/A
C_reopdisl	0	0	0	0.64	0.38	2.98	0	0	0.04	0.04	7.27	N/A
C_dvt	0	0	0	0.03	0	1.95	0.49	0	0	0	5.27	N/A
C_pe	0	0	0	2.31	1.72	10.02	3.06	0	0.02	0.02	4.24	N/A
C_reopinf	0	0	0	2.33	1.68	6.9	0.1	0	0.01	0.01	6.86	N/A
C_rev	0	0	0	2.99	2.21	10.79	3.36	0	0.03	0.03	9.7	N/A
RR_dvt	0	0	0	1.4	0.32	0.1	1.19	0	0.01	0.01	6.93	N/A
RR_inf	0	0	0	3.13	1.59	0.15	0.85	0	0.01	0.01	5.47	N/A
RR_pe	0	0	0	3.34	1.92	0.15	5.09	0	0.03	0.03	9.24	N/A
RR_omort	0	0	0	2.05	0.68	0.17	2.94	0	0.08	8.67	0	N/A

CV=Contribution to variance; DM= dominance measure; EC=elasticity coefficient ; EVPPI= expected value of partial perfect information; MCS= Monte Carlo simulation; MI=mutual information; MSD=maximum separation distance; RC=rank correlation; SB=sequential bifurcation; SRC=standardized regression coefficient

4.12. Correlation Analysis

Standard correlation test on the calibrated scores and the ranks was conducted to determine the correlation among the different methods. Results are presented in the table 26 and table 27 below.

The three EVPPI methods had a high degree of correlation with each other. The two methods that had the highest correlation with all three methods of estimating EVPPI (correlation coefficient (ρ) >0.98 for EVPPI- double MCS method, >.87 for EVPPI-quadrature and > 0.94 for single EVPPI) were contribution to variance and dominance measure methods. Mutual information, rank correlation, standardized regression coefficients, ANCOVA and maximum separation distance also had strong correlation with EVPPI ($\rho \geq 0.80$ for all EVPPI methods). Elasticity coefficients and sequential bifurcation had the lowest correlation coefficients with EVPPI results. EVPPI single MCS and EVPPI quadrature methods were highly correlated with EVPPI double MCS method with correlation scores >0.9. Among the other methods, contribution to variance was highly correlated to mutual information methods, rank correlation, dominance measure and ANCOVA. The highest correlation coefficient (0.999) was observed among ANCOVA and rank correlation.

If numerical values of the measures of importance are disregarded and ranking of the parameters is solely focus of interest, than the correlation analysis reveals that EVPPI double MCS method and EVPPI quadrature method have correlation factor 1, as well as ANCOVA and dominance measure method, since they resulted with identical ranking. EVPPI single method resulted with high correlation score with EVPPI quadrature and EVPPI double MCS (0.973), as well as contribution to variance with mutual information (0.906). The rest of the methods were less correlated (<0.85).

Table 26: Correlation Analysis among Calibrated Scores of the Results of Importance Measures

	EVPPI single	EVPPI quad	EVPPI double	CV	MI	EC	RC	DM	ANCOVA	SRC	MSD	SB
EVPPI single	1.000											
EVPPI quad	0.959	1.000										
EVPPI double	0.979	0.901	1.000									
CV	0.944	0.872	0.981	1.000								
MI	0.932	0.865	0.967	0.995	1.000							
EC	0.700	0.759	0.688	0.724	0.770	1.000						
RC	0.877	0.809	0.937	0.981	0.983	0.743	1.000					
DM	0.952	0.886	0.983	0.999	0.995	0.738	0.979	1.000				
ANCOVA	0.871	0.800	0.936	0.980	0.981	0.739	0.999	0.978	1.000			
SRC	0.871	0.800	0.935	0.979	0.980	0.737	0.998	0.977	0.999	1.000		
MSD	0.884	0.814	0.942	0.983	0.986	0.751	0.996	0.981	0.996	0.992	1.000	

CV=Contribution to variance, DM=dominance measure EC=elasticity ,MI=mutual information, MSD=maximum separation distance, RC=rank correlation, SB=sequential bifurcation, SRC=standardized regression coefficient

Table 27: Correlation Analysis among Ranking of the Results of Importance Measures

	EVPPI single	EVPPI quad	EVPPI double	CV	MI	EC	RC	DM	ANCOVA	SRC	MSD	SB
EVPPI single	1.000											
EVPPI quad	0.973	1.000										
EVPPI double	0.973	1.000	1.000									
CV	0.626	0.693	0.693	1.000								
MI	0.681	0.697	0.697	0.906	1.000							
EC	0.308	0.332	0.332	0.428	0.380	1.000						
RC	0.583	0.692	0.692	0.575	0.496	0.318	1.000					
DM	0.580	0.688	0.688	0.632	0.550	0.499	0.710	1.000				
ANCOVA	0.580	0.688	0.688	0.632	0.550	0.499	0.710	1.000	1.000			
SRC	0.727	0.739	0.739	0.653	0.710	0.387	0.587	0.846	0.846	1.000		
MSD	0.526	0.643	0.643	0.604	0.499	0.401	0.681	0.691	0.691	0.525	1.000	
SB	0.610	0.681	0.681	0.571	0.586	0.194	0.570	0.567	0.567	0.614	0.532	1.000

CV=Contribution to variance, DM=dominance measure; EC=elasticity ,MI=mutual information, MSD=maximum separation distance, RC=rank correlation,SB=sequential bifurcation, SRC=standardized regression coefficient

4.13. Computational Performance

The computational performance of the twelve implemented methods was measured by comparing the run time for each of the methods when being implemented to the case study MI THR versus STD THR. All calculations were done on a computer running Windows version 7 on 64-bit operating system, Intel Core i3-2377M processor at 1.50GHz and 7.8GB RAM.

All methods were performed in MS Excel (version 2007, Microsoft Corporation, Redmond, WA, USA), with only 2 techniques requiring additional add-ons in Excel, i.e. XLSTAT for ANCOVA and OLS Regression and OLSReg Function for regression coefficient method.

The table 30 below summarizes the number of runs required for each of the techniques, while table 31 includes the run time for each of the techniques. The recorded time represents the run time associated with running the necessary number of simulations and it does not include time spent on setting up algorithm.

Table 30: Number of runs required for each of the methods

Methods	Number of runs required to measure importance of single parameter	Number of runs required to measure importance of all parameters
EVPPI-double MCS	r*r	r*r*k
EVPPI-quadrature	r*m	r*m*k
EVPPI-single MCS	r	r*k
CV	r*m	r*m*k
MI	r*m	r*m*k
EC	r*m	r*m*k
RC	r	r
DM	r	r
ANCOVA	r	r
SRC	r	r
MSD	r	r
SB^	N/A	N/A
k = uncertain parameters (for case study k =26) r = replications within a MCS (for case study r =5000) m = sample values for input parameter of interest for quadrature (for case study m=101)		
CV=Contribution to variance, DM=dominance measure; EC=elasticity ,MI=mutual information, MSD=maximum separation distance, RC=rank correlation, SB=sequential bifurcation, SRC=standardized regression coefficient		

^ SB method is not Bayesian by nature and therefore number of runs required as a measure is inapplicable to it.

Table 31: Computational runtime associated with each of the methods

Methods	Runtime per parameter (minutes)	Total Runtime for all parameters (minutes)
EVPPI-double MCS	701	18,226
EVPPI-quadrature	149	3,874
EVPPI- single MCS	36	936
CV	149	3,874
MI	149	3,874
EC	149	3,874
RC	36	36
DM	36	36
ANCOVA	36	36
SRC	36	36
MSD	36	36
SB	97	97

CV=Contribution to variance, DM=dominance measure; EC=elasticity ,MI=mutual information, MSD=maximum separation distance, RC=rank correlation, SB=sequential bifurcation, SRC=standardized regression coefficient

Based on the results of the correlation analysis, contribution to variance and dominance measure were the two methods with the highest correlation scores with EVPPI. Based on the results in Table 31, the dominance measure is associated with much shorter runtime than the contribution to variance and therefore, dominance measure emerged as the best candidate for screening technique.

Based on the runtime results produced in the case study (Table 31), performing EVPPI double-MCS for all 26 input parameters in the model would result with 18,226 minutes

(12 days and 14 hours). Implementing the dominance measure as a screening technique would require a runtime of 36 minutes. Using the results of the dominance measure method as a screening technique and only performing EVPPI double MCS to 6 parameters identified as potentially important would result with a run time of 4,206 minutes, i.e. with a total runtime of 4,242 minutes (2 days and 22 hours).

Table 28: Decrease in runtime by implementing dominance measure as a screening technique

Methods	Runtime (minutes)	Decrease in runtime (%)
EVPPI-double MCS	18,226	
DM screening + EVPPI-double MCS	4,242	77%

Chapter 5

Discussion and Conclusion

5.1 Discussion

EVPPI is the current standard and has been presented as the theoretically correct method for determining parameter importance when the objective is to determine which parameters may be candidates for further research (Coyle 2004). The main advantages of EVPPI as a measure of importance are that:

1. it combines both uncertainty of the parameter and its importance
2. it provides evidence of whether uncertainty matters
3. it quantifies the monetary impact of making wrong decision
4. It does not require linear relationship among input parameters and the output

(Claxton 2008).

EVPPI is consistent with both the maximization of expected value and the economic concept of marginal reasoning (Felli and Hazen1998).

EVPPI does place an upper bound on the value of additional evidence and, hence, provides a screening mechanism whereby EVPPI needs to be ‘large enough’ to justify potential further research (Claxton 2008, Eckermann 2010). However, as Eckermann noted, in making this argument, a critical issue arises in relation to how EVPPI is determined to be ‘large enough’ or ‘not large enough’. To establish a threshold value for EVPPI requires consideration of the costs of undertaking research, which, in turn, is dependent on the proposed research type and size (Eckermann et al 2010). Furthermore, Eckermann argues that a sound value of information measure needs to address the following four questions:

1. Is further research for a health technology assessment (HTA) potentially worthwhile?
2. Is the cost of a given research design less than its expected value?
3. What is the optimal research design for an HTA?
4. How can research funding be best prioritized across alternative HTAs (Eckermann et al 2010)?

Using Occam’s razor approach, Eckermann concludes EVPPI provides neither a necessary nor a sufficient condition to address question 1, given that what EVPPI needs to exceed varies with the cost of research design.

Evaluating EVPPI does come at a computational cost, since the calculation of EVPPI may require many simulations runs of the same probabilistic model, which can become large burden given the time constraints the decision makers often face. Therefore, for the purpose of this thesis, alternative measures of importance were implemented and assessed and their results were compared with EVPPI results. The appropriateness of the other measures was dependant on similarity of the results with EVPPI results. Computational complexity, as well as means of interpreting results were assessed and factored in with an objective to come up with a recommendation for a screening method that will identify subgroup of parameters with potentially high EVPPI. That way, estimating EVVPI can be reserved only for those parameters, which will result with overall decrease in run-time.

Estimating EVPPI involves both an inner and an outer expectation. Three different methods for estimating EVPPI were implemented which vary by the methods employed to estimate these integrations- EVPPI double MCS, EVPPI quadrature and EVPPI single MCS. . The EVPPI quadrature method and the EVPPI double MCS method had identical ranking . However the EVPPI quadrature method may be preferred as it was associated with a much shorter run time. The quantitative values of EVPPI based on the quadrature method were lower than EVPPIs obtained through the double MCS methods, which might be the result of the potentially greater MCS maxima bias associated with the double MCs method as the inner integral is estimated through MCS. This outcome is of special interest, since EVPPI double-MCS method is computationally highly inefficient method requiring large number of iterations. EVPPI results identified cost of MI and STD THR as the most important parameters, followed by incremental utilities at 6, 12 and 3 months. Relative risk of revision scored 6^{th}, and the rest of parameters resulted with zero EVPPI.

Alternative measures of parameter importance can be assessed by how closely they replicated the results based on EVPPI, their ability to act as screening method and their computational complexity.

Excluding the elasticity-based technique and sequential bifurcation, the rest of the techniques identified the same 6 parameters as most important ones, although the ranking varied slightly. Thus we can exclude elasticity-based technique and sequential bifurcation as potential measures.

To be useful as a screening method, the measures must meaningfully identify parameters with potentially high EVPPIs. To assess this it is necessary to consider how these measures of importance can be interpreted. If we adopt the definition of importance to be the probability of making an incorrect decision and the consequences of such a decision, then measures of importance that focus solely on the degree of dispersion of the outcome measure may be inappropriate. As discussed before, one could have a high degree of correlation between a parameter and the outcome but there may be minimal risk of making a wrong decision. This suggests that for certain parameters threshold values would be required to be adopted to determine which parameters are important and which are not.

Interestingly, EVPPI methods resulted with non-zero EVPPI only for 6 out of 26 parameters. A similar pattern was observed with dominance measure. Therefore, in this case study we didn't have to decide on threshold on importance. We could simply argue that only the parameters with non-zero EVPPI are important, and the rest are not. However, the rest of the methods resulted with non-zero values for all of the parameters, and therefore the need of setting up a threshold is inevitable, which indirectly implies incorporating subjectivity to these methodologies. Thus, with respect to this criterion the dominance measure would be the preferred screening method.

Techniques differ in terms of computational complexity. Some techniques are relatively simple requiring only the analysis of a single data base obtained with 5000 runs Monte Carlo simulation, such as rank correlation, dominance measure, ANCOVA, standardized regression coefficient and maximum separation distance. However, other techniques require multiple repeat simulations.

For complex models where there are a significant number of uncertain input parameters, screening to identify parameters which should be subject to more complex analysis of uncertainty may be desirable. Therefore, the role of the alternative importance measures is to act as a screening method to identify a subset of parameters which would go under the time-consuming EVPPI algorithm. Since the main purpose of this approach would be reducing run time, the methods that are computationally as efficient as EVPPIs would be eliminated. Therefore, contribution to variance, elasticity and mutual information methods are not candidates for screening method.

Based on computational complexity and ranking, dominance measure, standardized regression coefficient, rank correlation, ANCOVA and maximum separation distance are 5 methods candidates to be successful as screening methods. Dominance measure resulted with very highly correlated calibrated scores with EVPPI calibrated scores and had the same distribution of positive and zero values as the EVPPI methods. Thus, it might be the best alternative measure for screening parameters of importance. Furthermore, based on the runtime calculations obtained from the case study, implementing dominance measure as a screening technique would result with 77% reduction in runtime. (Table 32)

It should be noted that the health economic model used in this research as a case study is of relatively low complexity. The model included a total of 26 input parameters and 7 health states. There are no time-dependant parameters, the cycle length is 1 year and the total number of cycles is 40. Also, there is no correlation incorporated among the input parameters.

Often health economic models with higher analytical complexity than the case study are constructed. Namely, health economic models often include higher number of health states and significantly higher numbers of input parameters. The cycle length could also be shorter which would result with higher total number of cycles. Also, health economic models could incorporate time-dependant variables and correlation among variable could be modelled.

The analytical complexity of the model is not expected to have an impact to the accuracy of the importance measure results; nevertheless the impact of the model complexity to the runtime is expected to be high and the positive impact of performing screening method is expected to be even more pronounced when implemented in more complex methods.

5.2 Conclusion

EVPPI is the current gold standard measure of importance in health economic evaluations when the focus is in identifying parameters for which further research may be justified.

The complexity of performing EVPPI as well as the long run time suggests the need to identify alternative methods for identifying parameters. Moreover, with the current shift in reimbursement policies towards risk sharing agreements and value-based pricing of health technologies, performing importance analysis and determining the expected value of further research will become much more pertinent.

Findings of this research suggested that prior to performing value of information analysis and calculating EVPPI, one can perform a screening method to identify subgroup of parameters as candidates for being most important parameters. Successful screening method would significantly reduce the time and effort of determining most important parameters. Based on this research, the dominance measure method emerged as the most favourable screening method that could be applied prior to conducting EVPPI analysis. Instead of determining EVPPI for each parameter in the model, this thesis suggest that the dominance measure method can be used to identify those parameters that can potentially have significantly high EVPPI and therefore the complex methods for estimating EVVPI can be reserved for those parameters.

Given the limited resources available for health care, the value of performing importance analysis within health economic evaluation is evident. The opportunity cost of not performing it can be considerable. Importance analysis needs to be an indispensable step in evaluating new health technologies and a standard practice for such evaluations. Advancing the practical methods for conducting importance analysis and developing more accessible methods is highly desirable. Until then, applying a screening method, such as dominance measure technique before conducting an EVPPI analysis is an efficient option to determine the most important parameters in terms of contribution to decision uncertainty.

REFERENCES:

Azen, R., & Budescu, D. V. (2003). The dominance analysis approach for comparing predictors in multiple regression. Psychological methods, 8(2), 129-31

Eriksson, B. I., Borris, L. C., Friedman, R. J., Haas, S., Huisman, M. V., Kakkar, A. K., ... & Geerts, W. (2008). Rivaroxaban versus enoxaparin for thromboprophylaxis after hip arthroplasty. New England Journal of Medicine, 358(26), 2765-2775

Bettonvil, B,., and Kleijnen J. P. C.. (1997). Searching for important factors in simulation models with many factors: Sequential bifurcation. European Journal of Operational Research 96 (1), 180-194

Birch S, Gafni A, (2006). Information created to evade reality (ICER): things we should not look to for answers, Pharmacoeconomics.24(11), 1121-31

Blalock H. M. (1963). Jr., Correlated Independent Variables: The Problem of Multicollinearity Social Forces (1963) 42 (2), 233-237

Borgonovo, E. (2006). Measuring uncertainty importance: investigation and comparison of alternative approaches, Risk Analysis, 26(5), 1349-1361.

Brazier J, Roberts J, Deverill M. (2002). The estimation of a preference based measure of health from the SF36. J Health Econ, 21, 271-92

Brennan A , Kharroubi S , O'Hagan A Brennan et Al. (2007). Calculating Partial Expected Value Of Perfect Information in Cost-Effectiveness Models, Medical Decision Making, 27 (4), 448-470.

Brennan A., Chick S. E., Davies R. (2006) Taxonomy of Model Structures for Economic Evaluation of Health Technologies, Health Econ. 15(12), 1295-310

Briggs, A. (2005). Probabilistic Analysis of Cost-Effectiveness Models: Statistical Representation of Parameter Uncertainty. Value in Health, 8(1), 1-2

Briggs, A., & Sculpher, M. (1998). An introduction to Markov modelling for economic evaluation. Pharmacoeconomics, 13(4), 397-409

Briggs, A. H. (1999). A Bayesian approach to stochastic cost-effectiveness analysis. Health Economics, 8(3), 257-261

Briggs, A. H., Weinstein, M. C., Fenwick, E. A., Karnon, J., Sculpher, M. J., & Paltiel, A. D. (2012). Model Parameter Estimation and Uncertainty Analysis A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group–6. Medical Decision Making, 32(5), 722-732

Briggs, A. H. (2000). Handling uncertainty in cost-effectiveness models. Pharmacoeconomics, 17(5), 479-500

Bring, J. (1994). How to standardize regression coefficients. The American Statistician, 48(3), 209-213

Budescu D. V, (2003)The Dominance Analysis Approach for Comparing Predictors in Multiple Regression, Psychological Methods , 8(2), 129–148

Bujkiewicz S, Hayley E. Jones, Monica C.W. Lai, Nicola J. Cooper, Neil Hawkins, Hazel Squires, Keith R. Abrams, David J. Spiegelhalter, Alex J. Sutton, (2011) Development of a Transparent Interactive Decision Interrogator to Facilitate the Decision-Making Process in Health Care, Value in Health, 14(5), 768-776

Caro J, Briggs, A. H, Siebert U, Kuntz, K. M. (2012). Modeling Good Research Practices—Overview A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force–1. Medical Decision Making, 32(5), 667-677

Centre for Excellence for Joint Replacement: The Norwegian Arthroplasty Register, The Norwegian Cruiate Ligament Register, The Norwegian Hip Fracture Register. Bergen: Haukeland University Hospital: Department of Orthopedic Surgery; 2007. Available: http://www.haukeland.no/nrl/Report2006.pdf

Charles MN, Fejbel RJ, Kim P. (2006) Minimally invasive surgery of the hip - a randomized pilot study. Annual Meeting of the Canadian Orthopaedic Association; 2006; Toronto

Cheng R, (1997), Searching For Important Factors: Sequential Bifurcation Under Uncertainty, Proceedings of the 1997 Winter Simulation Conference

Claxton K, (1999) Bayesian approaches to the value of information: implications for the regulation of new pharmaceuticals, Health Econ. 8(3), 269-74

Claxton K, (2008) Exploring uncertainty in cost-effectiveness analysis, Pharmacoeconomics 26(9), 781-798

Claxton K.P.and Sculpher M.J., (2006), Using Value of Information Analysis to Prioritise Health Research Some Lessons from Recent UK Experience, Pharmacoeconomics, 24 (11), 1055-1068

Coyle D, Barbeau M, Baladi JF. (2001) Bayesian economic analysis of treatment with entacapone for Parkinson's Disease patients in Canada. Meeting of the Society of Medical Decision Making.

Coyle D, Buxton MJ, Bernie J.O.(2003) Measures of importance for economic analysis based on decision modeling. Journal of Clinical Epidemiology, 56(10), 989-997

Coyle D, Coyle K, Vale L, Verteuil R, Imamura M, Glazener C, Zhu S, CADTH Technology Report, HTA, (2008) Minimally Invasive Arthroplasty in the Management of Hip Arthritic Disease: Systematic Review and Economic Evaluation

Coyle D. (2004) The development and application of normative framework for considering uncertainty and variability in economic evaluation, PhD Thesis, Brunel University, 2004

de Kroon, H., Plaisier, A., van Groenendael, J., Caswell, H., (1986). Elasticity: the relative contribution of demographic parameters to population growth rate. Ecology 67, 1427– 1431

Detsky A. S, Naglie G, (1990). Clinician's Guide to Cost-Effectiveness Analysis, Annals of Internal Medicine, 113(2), 147-154

Drummond M. F, and Sculpher M.J. (2006). Better Analysis for Better DecisionsFacing Up to the Challenges, Pharmacoeconomics, 24 (11), 1039-1042

Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL, (2005). Methods for the economic evaluation of health care programmes, Third edition: Oxford: Oxford University Press; 2005

Eckermann S, Karnon and Willan J.R.(2010) , The Value of Value of Information, Best Informing Research Design and Prioritization Using Current Methods, Pharmacoeconomics 2010; 28 (9): 699-709

Felli, J. C., & Hazen, G. B. (1998). Sensitivity analysis and the expected value of perfect information. Medical Decision Making, 18(1), 95-109

Fenwick, E., Claxton, K., & Sculpher, M. (2008). The value of implementation and the value of information: combined and uneven development. Medical Decision Making, 28(1), 21-32.

Fieller, E. C., Hartley, H. O., & Pearson, E. S. (1957). Tests for rank correlation coefficients. I. Biometrika, 44(3/4), 470-481

Fox-Rushby J, and John Cairns J, (2005). Economic Evaluation, Division: Open University Press Book details, ISBN: 9780335218479

Frey HC, Patil SR., (2002). Identification and review of sensitivity analysis methods, Risk Anal. 2002 Jun; 22(3), 553-78

Gafni, A., & Birch, S. (2006). Incremental cost-effectiveness ratios (ICERs): The silence of the lambda. Social Science & Medicine, 62(9), 2091-2100

Greene, W.H., (2000). Econometric Analysis, Fourth Edition, Prentic-Hall, Upper Saddle River, New Jersey, 2000

Griffin S, Claxton K. I(2007). Interpreting the expected value of perfect information about parameters. Med Dec Making, 27 (1), 38 - E8

Hamby D.M (1994), A comparison of sensitivity analysis techniques, available online, reproduced from the journal Health Physics with permission from the Health Physics Society, 1994

Helton JC and Davis FJ. (2002). Illustration of sampling-based methods for uncertainty and sensitivity analysis. Risk Anal 22(3), 591–622

Helton, J. C. (1993), Uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive waste disposal. Reliability Engineering & System Safety, 42(2), 327-367.

Hofer, E. (1999). Sensitivity analysis in the context of uncertainty analysis for computationally intensive models. Computer Physics Communications, 117(1), 21-34

Hunter, C. M., Moller, H., & Fletcher, D. (2000). Parameter uncertainty and elasticity analyses of a population model: setting research priorities for Shearwaters. Ecological Modelling, 134(2), 299-324323

Iman R. L., Helton J.C, (1988), An Investigation of Uncertainty and Sensitivity Analysis, Techniques for Computer Models, Risk Analysis, 8(1), 71–90

Info Phys Comp'' Draft: November 9, 2007 http://www.stanford.edu/~montanar/RESEARCH/BOOK/partA.pdf

Lie, S. A., Havelin, L. I., Furnes, O. N., Engesaeter, L. B., & Vollset, S. E. (2004). Failure rates for 4762 revision total hip arthroplasties in the Norwegian Arthroplasty Register. Journal of Bone & Joint Surgery, British Volume, 86(4), 504-509

Life Tables, Statistics Canada, 2000 to 2002, Available: http://www.statcan.gc.ca/pub/84-537-x/4064441-eng.htm

Meltzer D, Basu A and Conti R, (2010), The Economics of Comparative Effectiveness Studies Societal and Private Perspectives and their Implications for Prioritizing Public Investments in Comparative Effectiveness Research, Pharmacoeconomics, 28 (10), 843-853

Meltzer, D. O., Hoomans, T., Chung, J. W., & Basu, A. (2011). Minimal modeling approaches to value of information analysis for health research. Medical Decision Making, 31(6), E1-E22.

Mokhtari, A., & Frey, H. C. (2006). Evaluation of sampling-based methods for sensitivity analysis: Case study for the E. coli food safety process risk model. Human and Ecological Risk Assessment, 12(6), 1128-1152

OCCI Costing Analysis Tool. Toronto: Ontario Case Costing Initiative; 2010 Available: http://www.occp.com

OCCI Costing Analysis Tool. Toronto: Ontario Case Costing Initiative; 2011 Available: http://www.occp.com

Ontario Ministry of Health and Long-term Care. (2010) Ontario Health Insurance (OHIP) Schedule of Benefit and Fees. Toronto, Government of Ontario

Parvizi, J., Johnson, B. G., Rowland, C., Ereth, M. H., & Lewallen, D. G. (2001). Thirty-day mortality after elective total hip arthroplasty. The Journal of Bone & Joint Surgery, 83(10), 1524-1528.

Roberts, M., Russell, L. B., Paltiel, A. D., Chambers, M., McEwan, P., & Krahn, M. (2012). Conceptualizing a Model A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force–2. Medical Decision Making, 32(5), 678-689.

Saltelli A, Chan K, Scott EM. (2000) Sensitivity Analysis. Wiley, Chichester
Saltelli, A., Tarantola, S., & Chan, K. S. (1999). A quantitative model-independent method for global sensitivity analysis of model output. Technometrics, 41(1), 39-56

Sculpher M and Claxton K, (2012), Real Economics Needs to Reflect Real Decisions, Pharmacoeconomics 2012; 30 (2): 133-136

Ihara, Shunsuke I(1993), Information theory for continuous systems. World Scientific. p. 2. ISBN 9789810209858

Smith, C.J., Croke, B.F.W. (2007) Uncertainty of Parameter Values in a Gully Erosion Model. Proceedings of the International Congress on Modeling and Simulation MODSIM2007 Modelling and Simulation Society of Australia and New Zealand Inc. (MSSANZ) , Oxley, L., Kulasiri, D. (Eds) , Christchurch, N.Z. , 2438-2444.

Sîeen, H., & Erikstad, K. F. (1996). Sensitivity of willow grouse Lagopus lagopus popuiatioi dynamics to variations m demographic parameters.

Sullivan PW, Ghushchyan V. (2006) Preference-Based EQ-5D index scores for chronic conditions in the United States. Med Decis Making, 26(4), 410-20

Tabachnick B.C, Fidell L. S, (1996), Using multivariate statistics, Book (ISBN 0673994147), 1996

Tappenden P, Chilcott JB, SEggington S, Oakley J and McCabe C, (2004), Methods for expected value of information analysis in complex health economic models:

developments on the health economics of interferon-β and glatiramer acetate for multiple sclerosis, Health Technology Assessment, 8(39)

Wan H., Ankenman B. E. and Nelson B. L., (2006) Controlled Sequential Bifurcation: A New Factor-Screening Method for Discrete-Event Simulation Operations Research, 54(4), 743 -755

Weinstein, M. C. (2006) Recent developments in decision-analytic modelling for economic evaluation, Pharmacoeconomics 24(11), 1043-1054

Winquist E, Bell CM, Clarke JT, Evans G, Martin J, Sabharwal M, Gadhok A, Stevenson H, Coyle D., (2012) An evaluation framework for funding drugs for rare diseases, Value Health, 15(6), 982-986

APPENDIX A

Generally, graphical methods are used to give visual indication of how an output is affected by variation in inputs (Frey and Patil 2002). Graphical methods can be used as a screening method before further analysis of a model or to represent complex dependencies between inputs and outputs and can be used to complement the results of mathematical and statistical methods for better representation (Frey and Patil 2002).

The following figures (Figure 21-46) represent scatter plots of the outcome (INB) versus the input parameters and give a visual impression of correlation between the input variables and the output. The figures are ordered by EVPPI of the input variables in a descending order. It is visible that the variables with non zero EVPPI have a more recognizable visual pattern in terms of interaction between the variable and INB.

Figure 21: Scatter Plot of Cost of MI THR versus INB

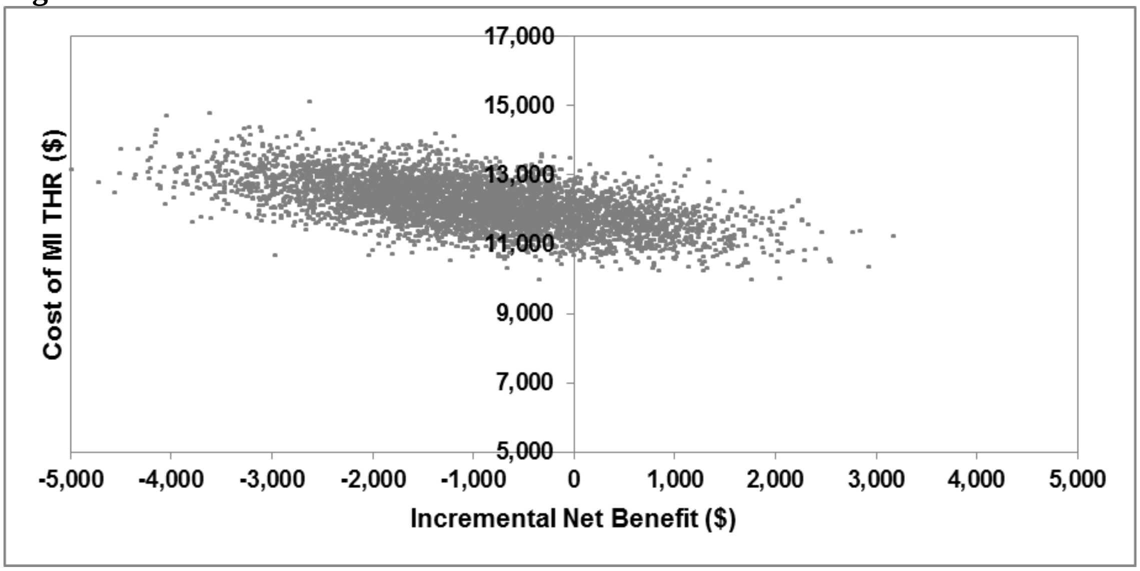


Figure 22: Scatter Plot of Cost of STD THR versus INB

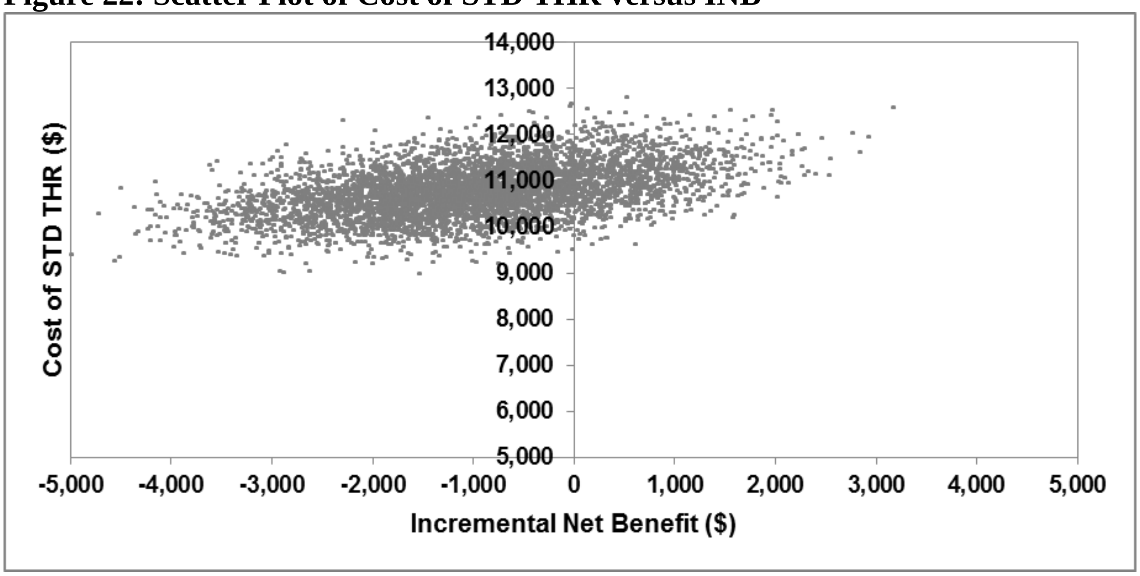


Figure 23: Scatter Plot of Incremental Utility at 6 months versus INB

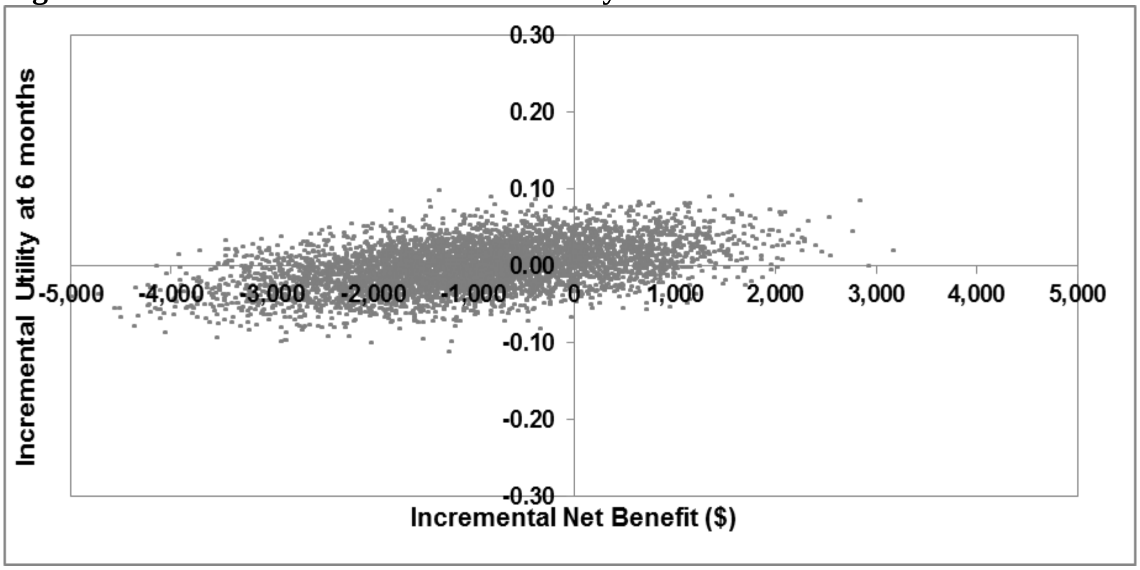


Figure 24: Scatter Plot of Incremental Utility at 12 months versus INB

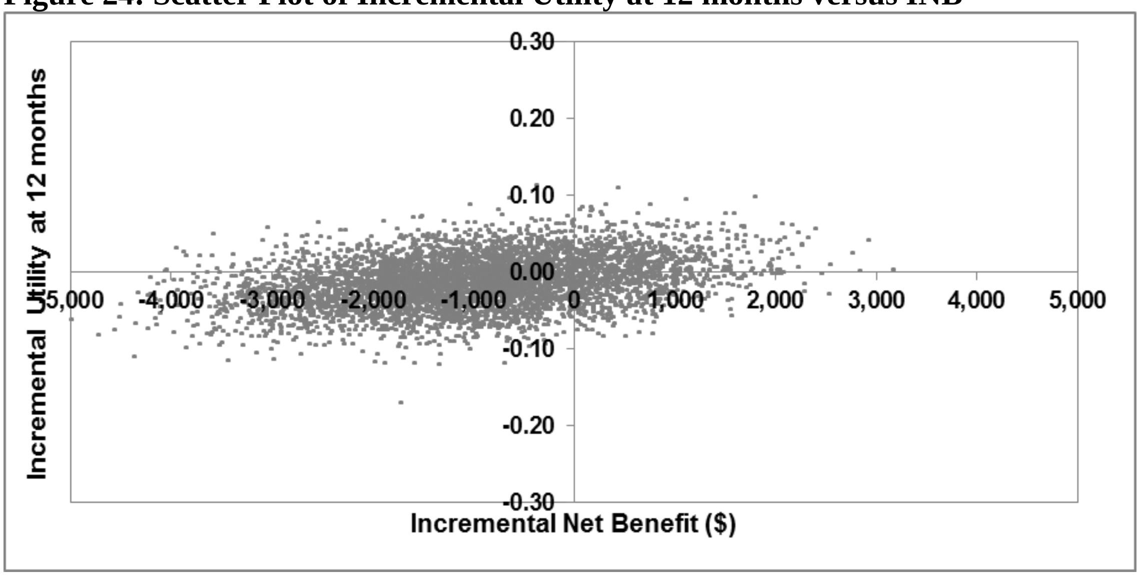


Figure 25: Scatter Plot of Incremental Utility at 3 months versus INB

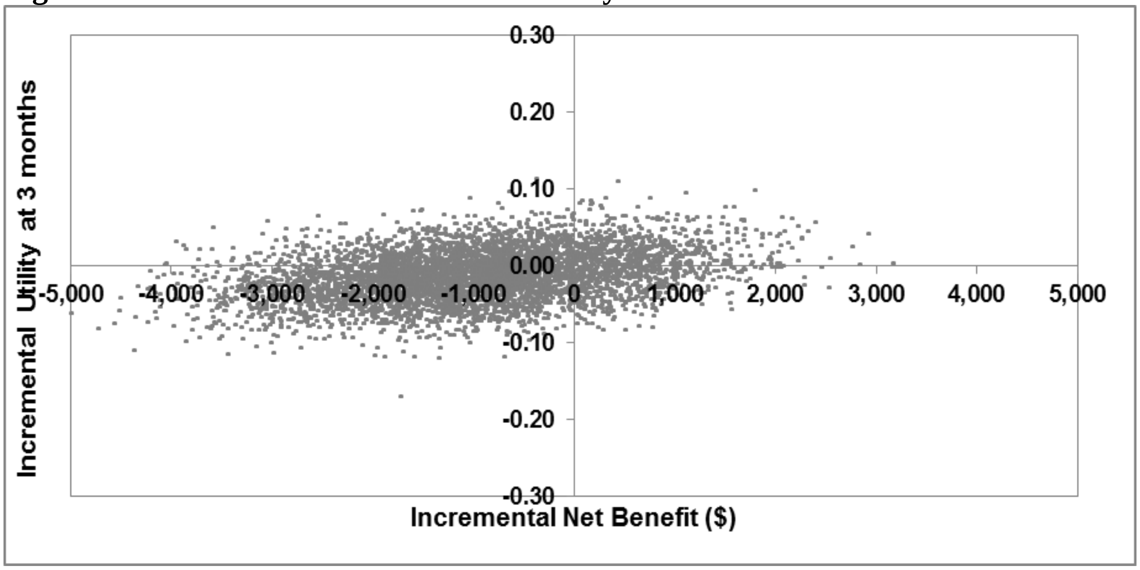


Figure 26: Scatter Plot of Relative Risk of Revision versus INB

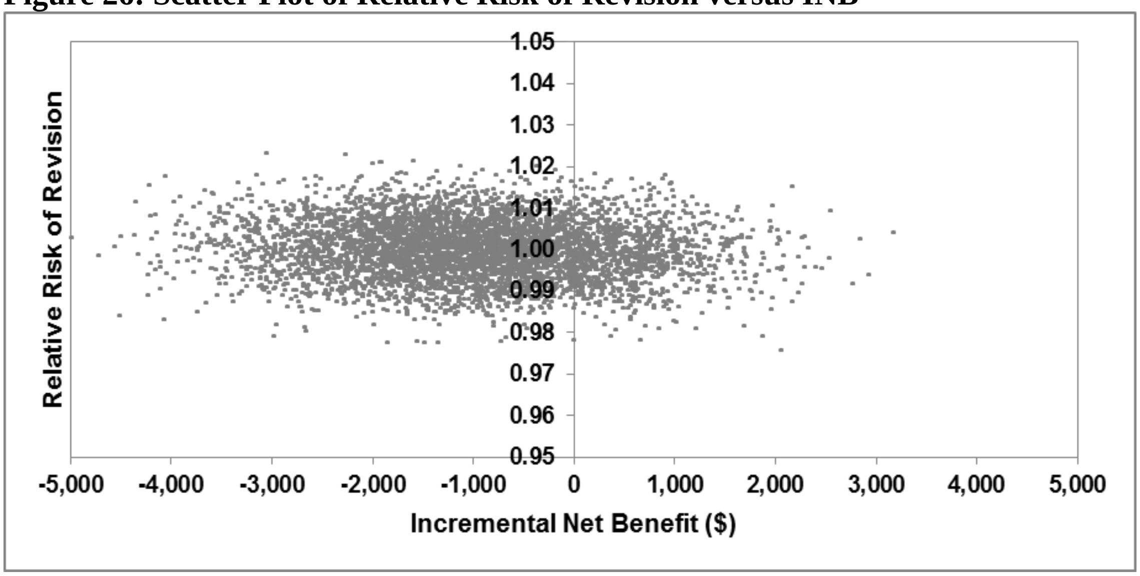


Figure 27: Scatter Plot of Cost of Reoperation due to Infection versus INB

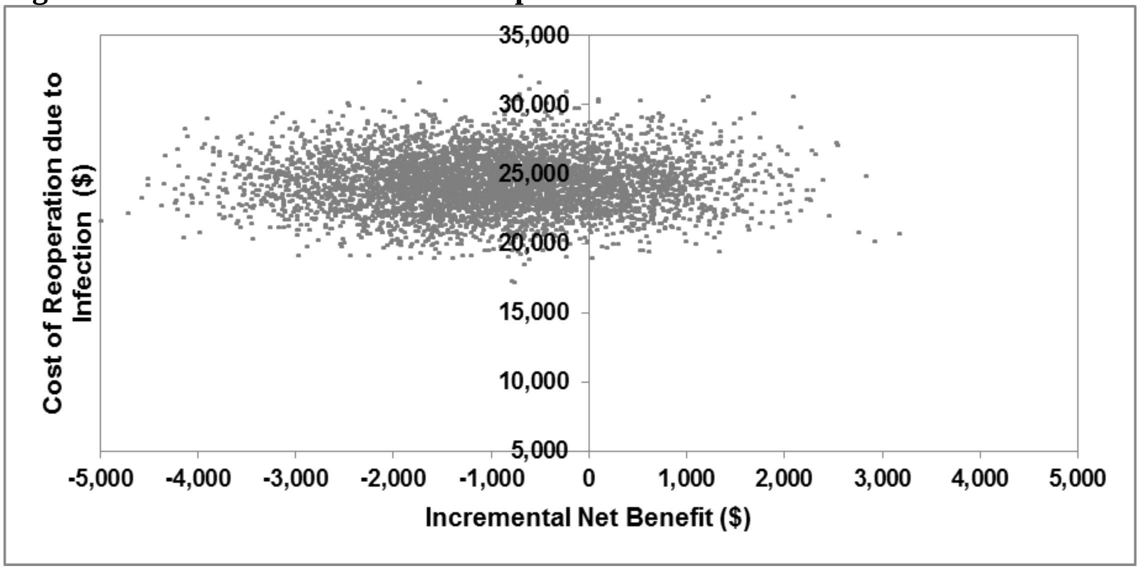


Figure 28: Scatter Plot of Cost of Revision versus INB

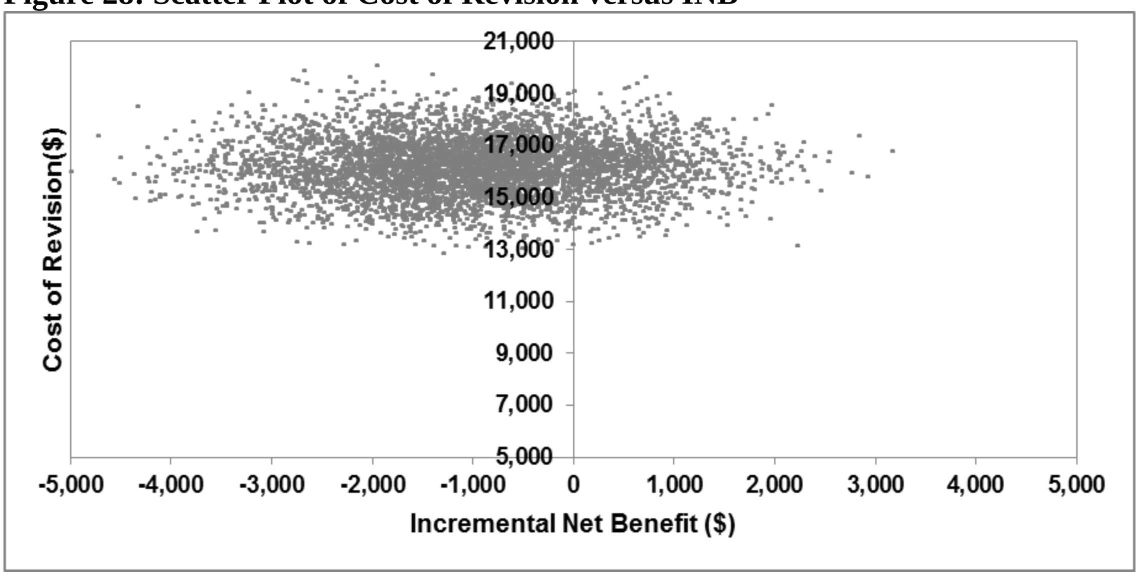


Figure 29: Scatter Plot of Cost of PE versus INB

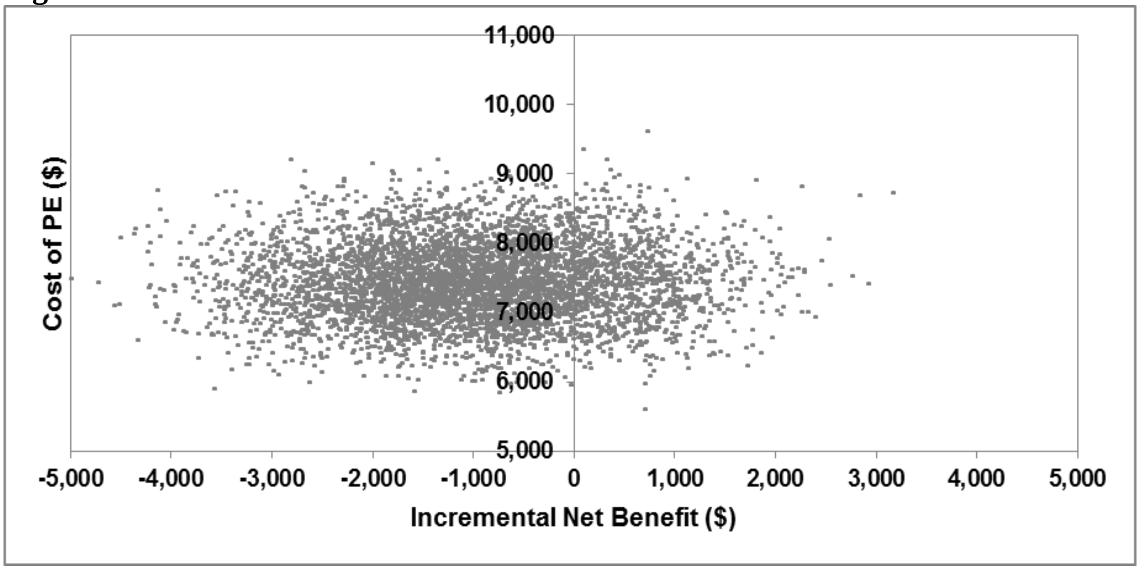


Figure 30: Scatter Plot of Cost of PE versus INB

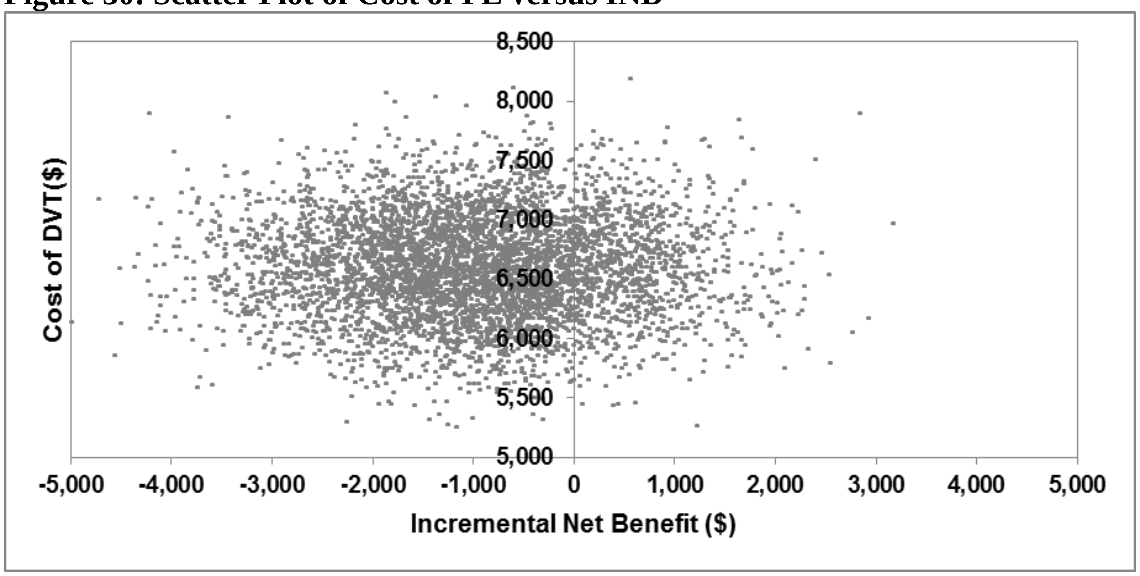


Figure 31: Scatter Plot of Cost of Dislocation versus INB

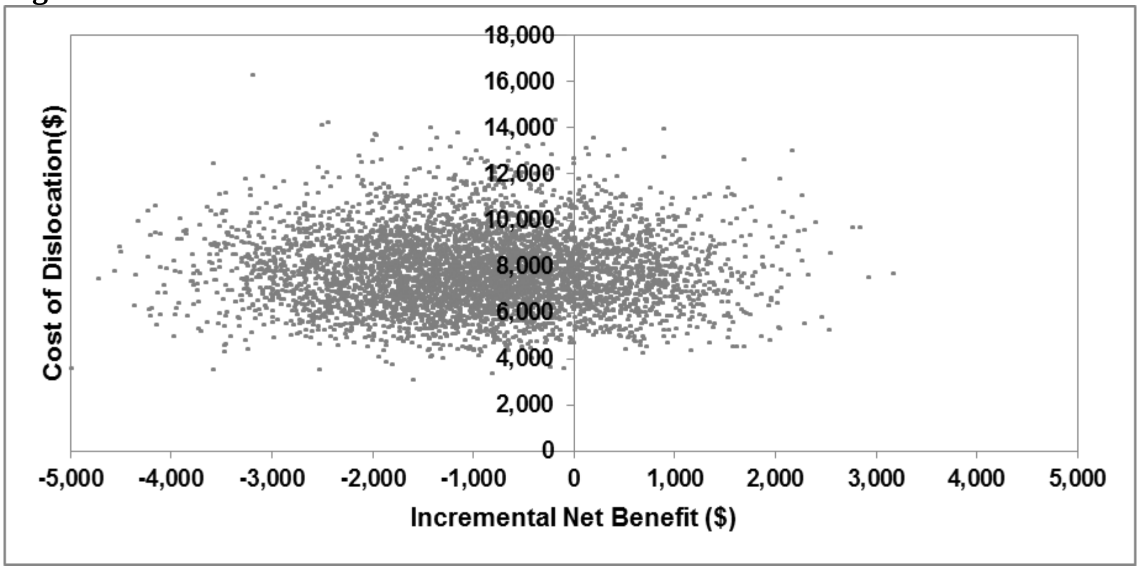


Figure 32: Scatter Plot of Relative Risk of Operative Mortality versus INB

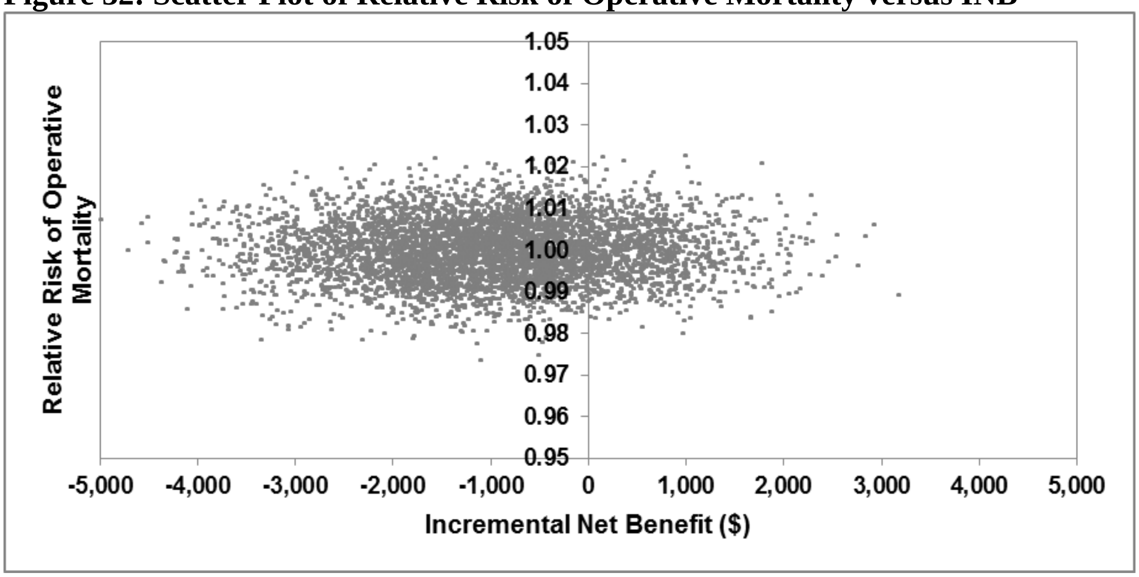


Figure 33: Scatter Plot of Relative Risk of PE versus INB

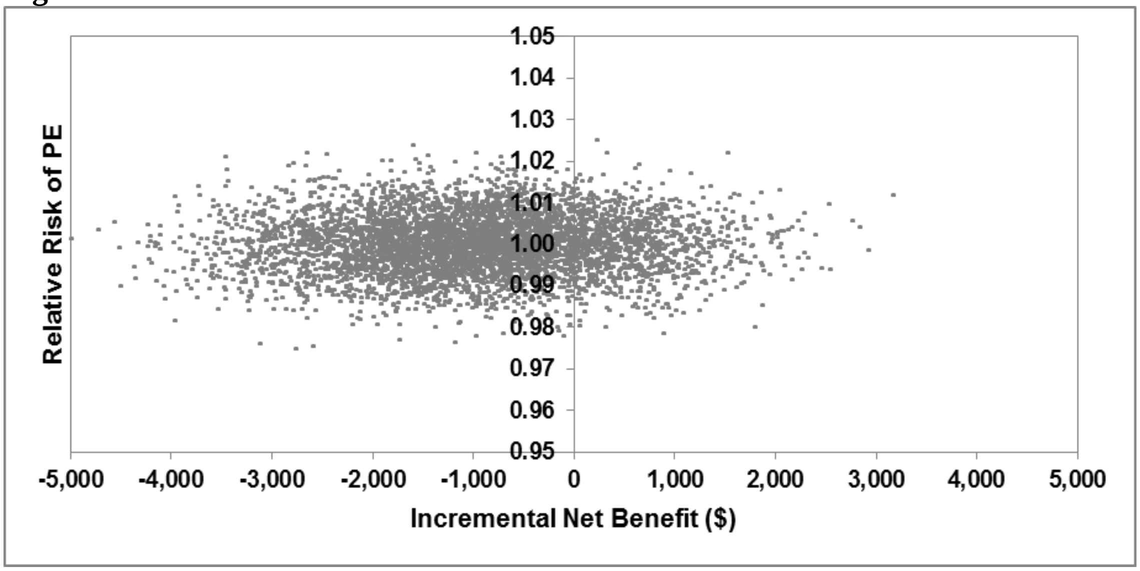


Figure 34: Scatter Plot of Relative Risk of Infection versus INB

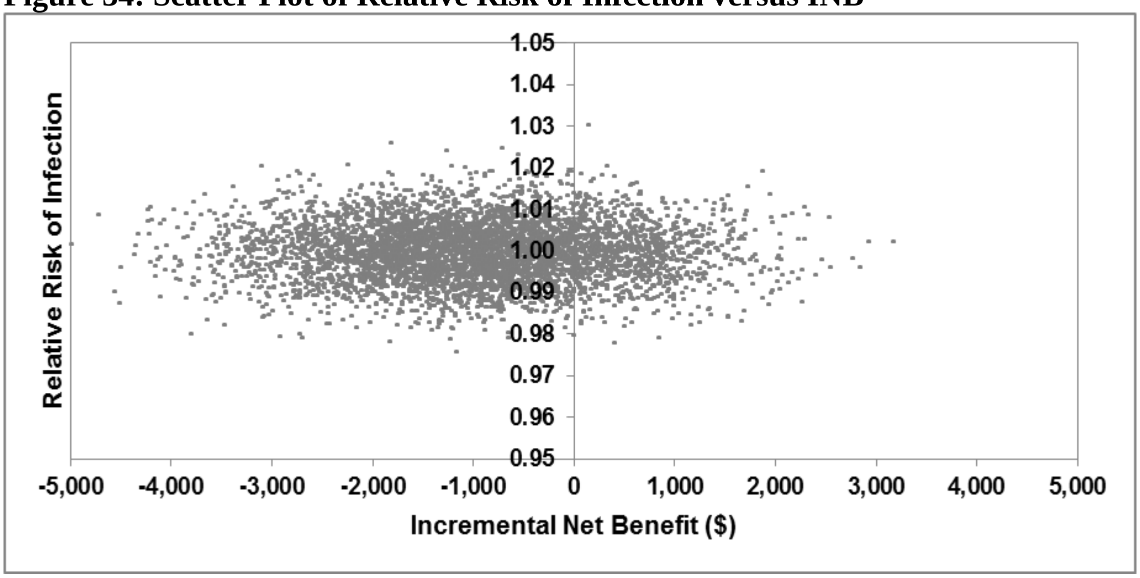


Figure 35: Scatter Plot of Relative Risk of DVT versus INB

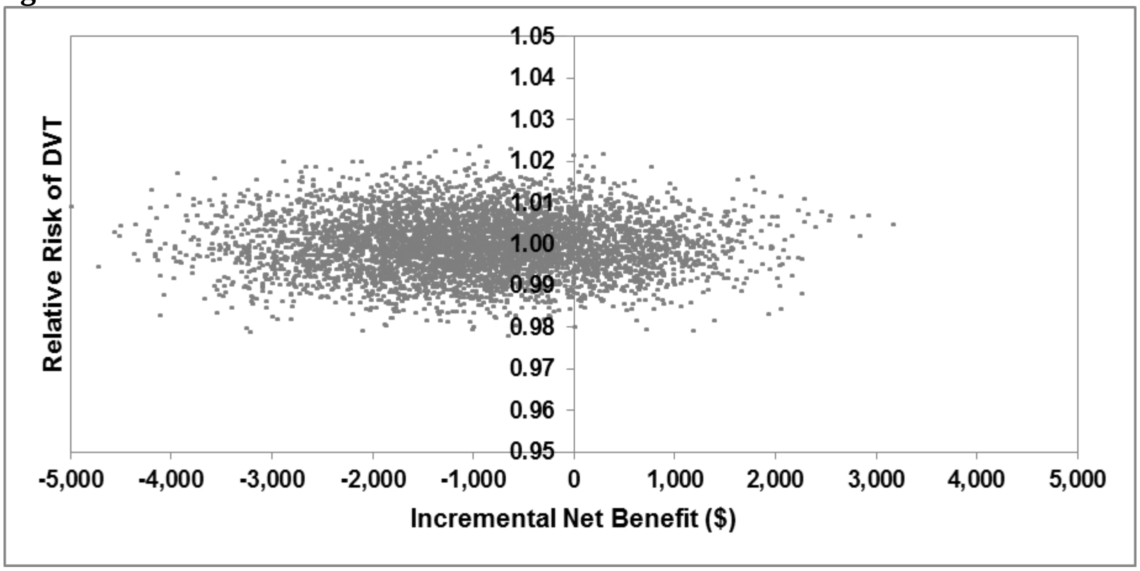


Figure 36: Scatter Plot of Utility at 3 months versus INB

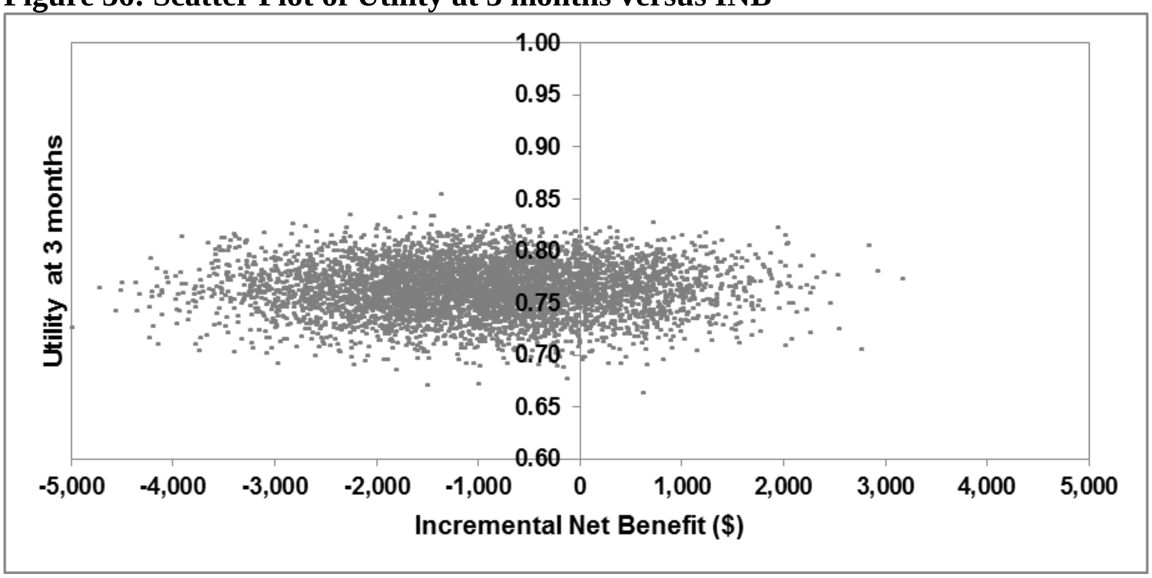


Figure 37: Scatter Plot of Utility at 6 months versus INB

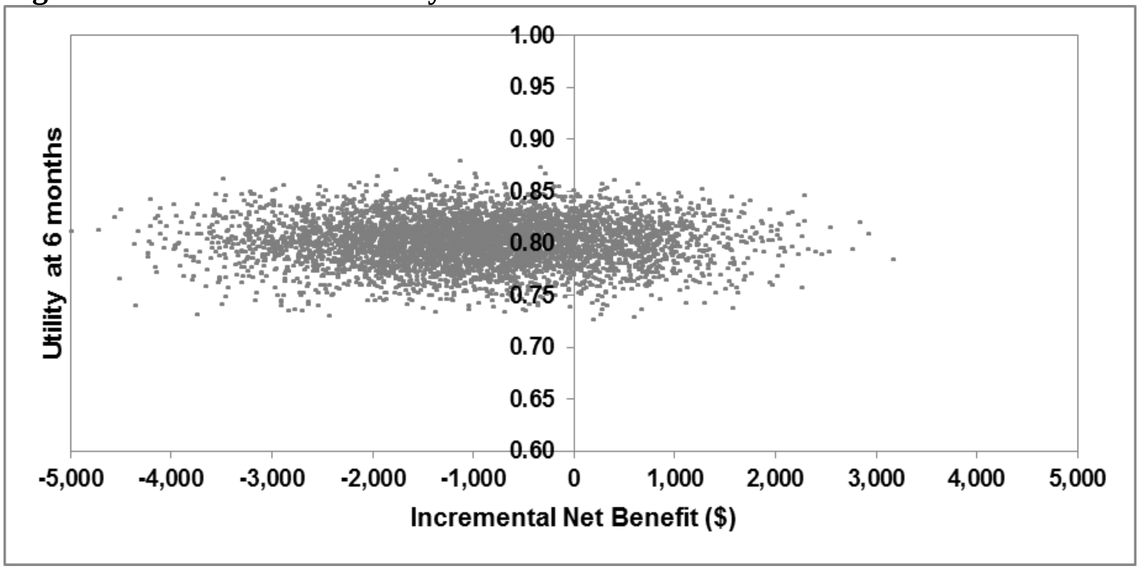


Figure 38: Scatter Plot of Utility at 12 months versus INB

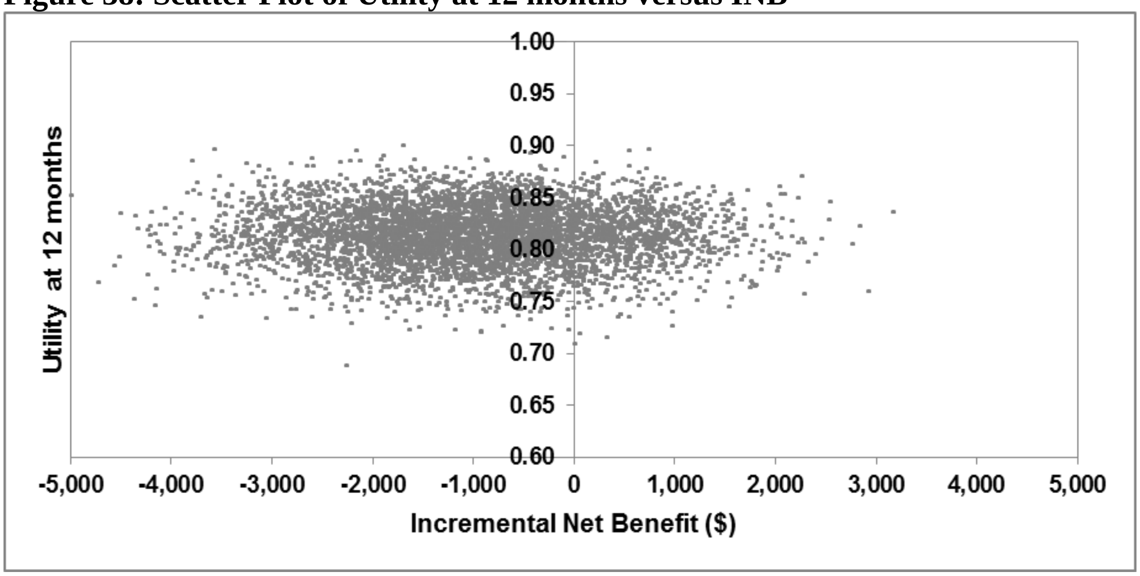


Figure 39: Scatter Plot of Preoperative Utility versus INB

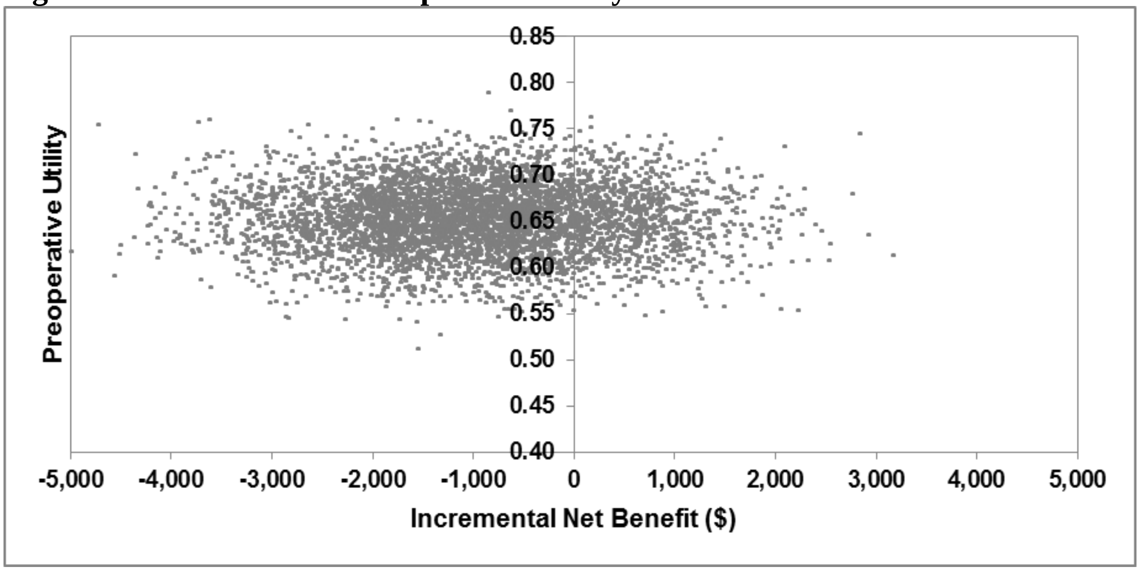


Figure 40: Scatter Plot of Probability of PE versus INB

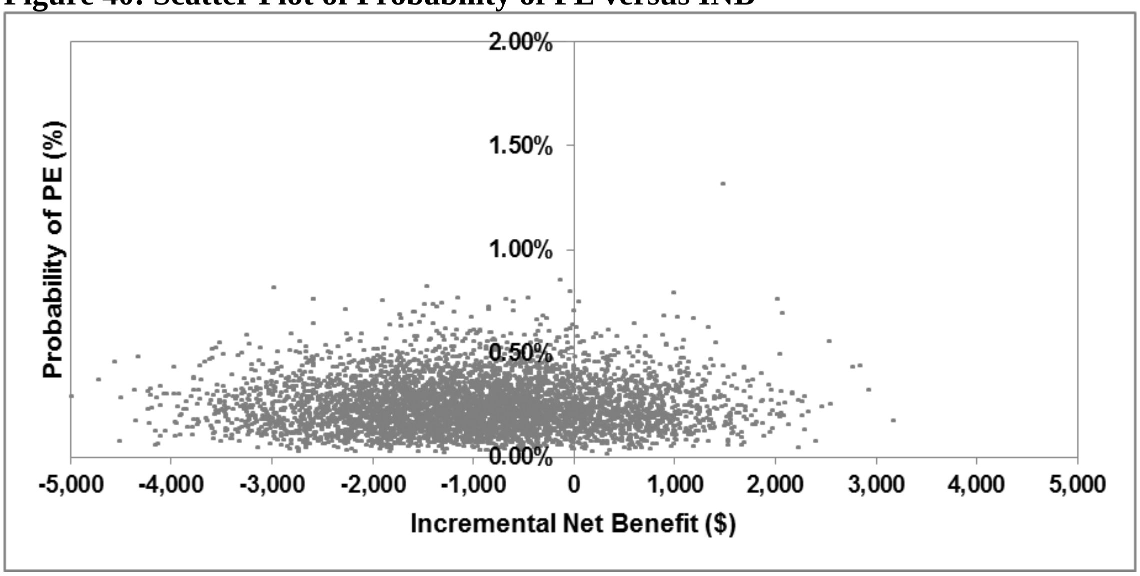


Figure 41: Scatter Plot of Probability of DVT versus INB

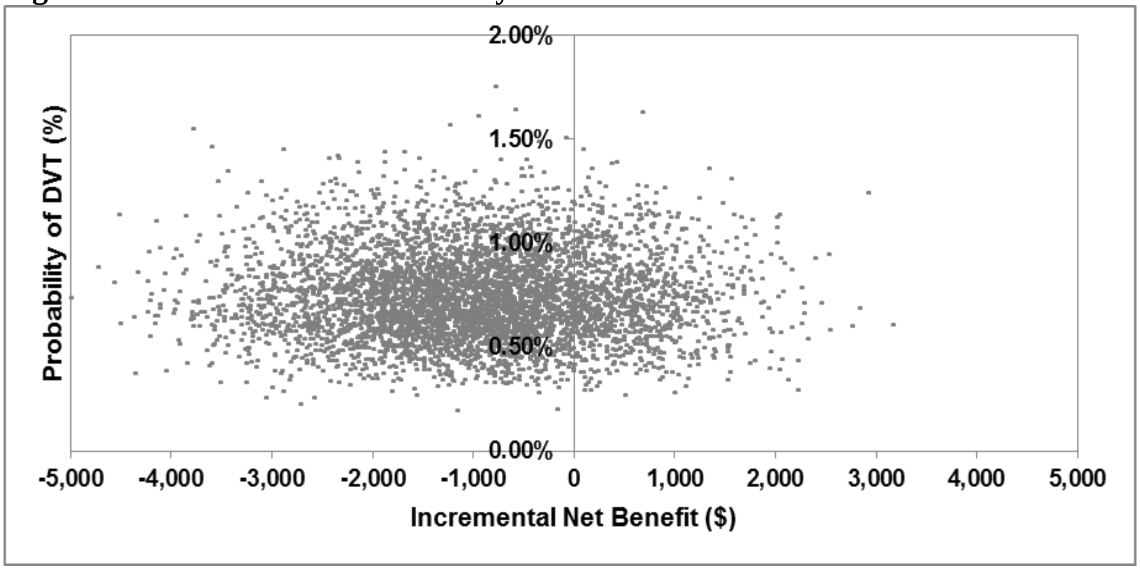


Figure 42: Scatter Plot of Probability of Operative Mortality versus INB

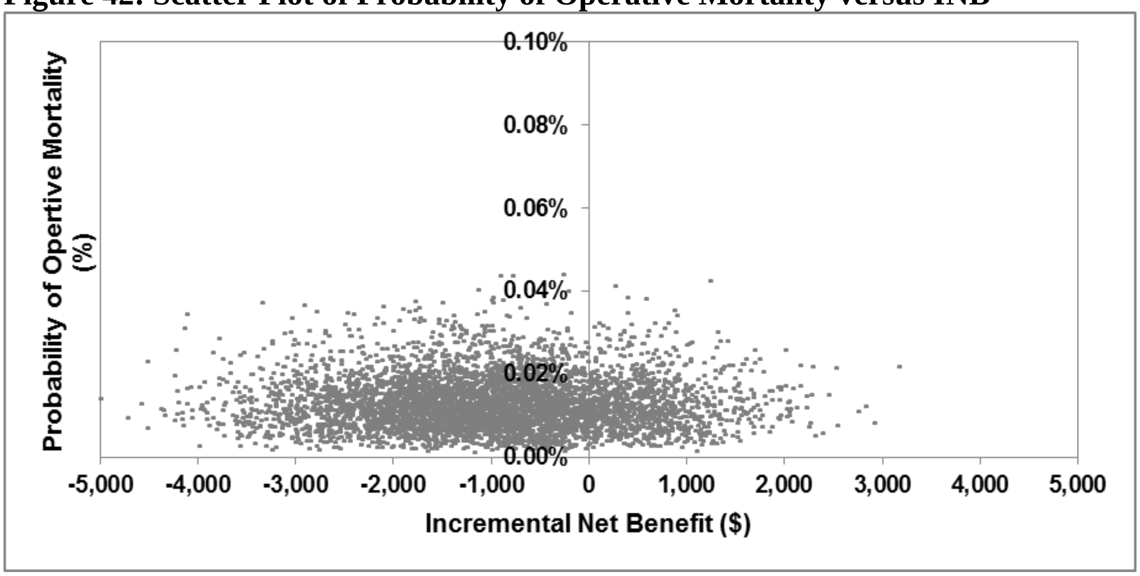


Figure 43: Scatter Plot of Probability of 1st Revision versus INB

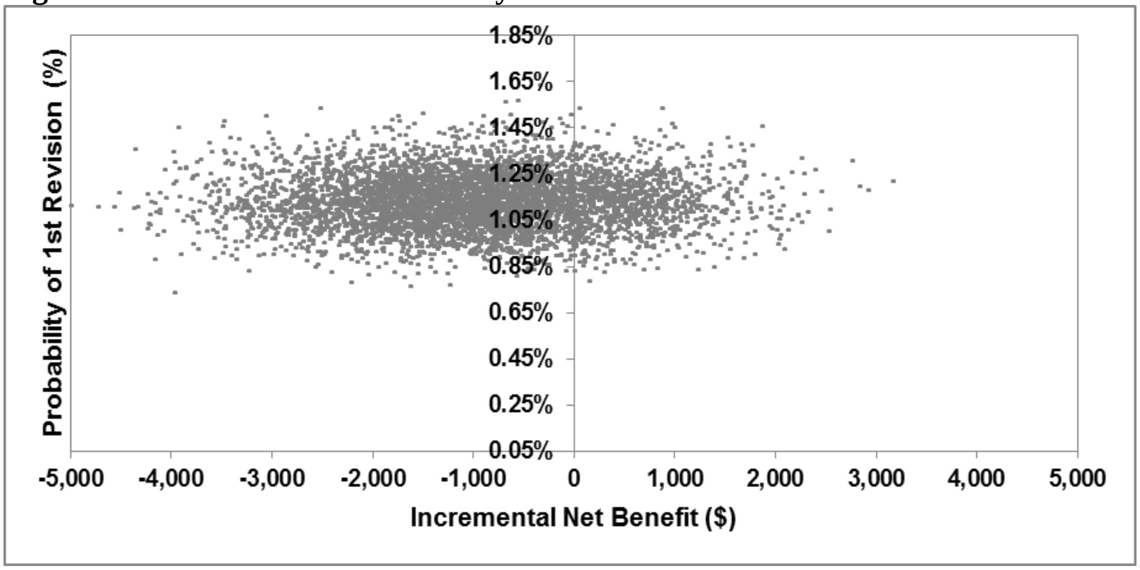


Figure 44: Scatter Plot of Probability of 2nd Revision versus INB

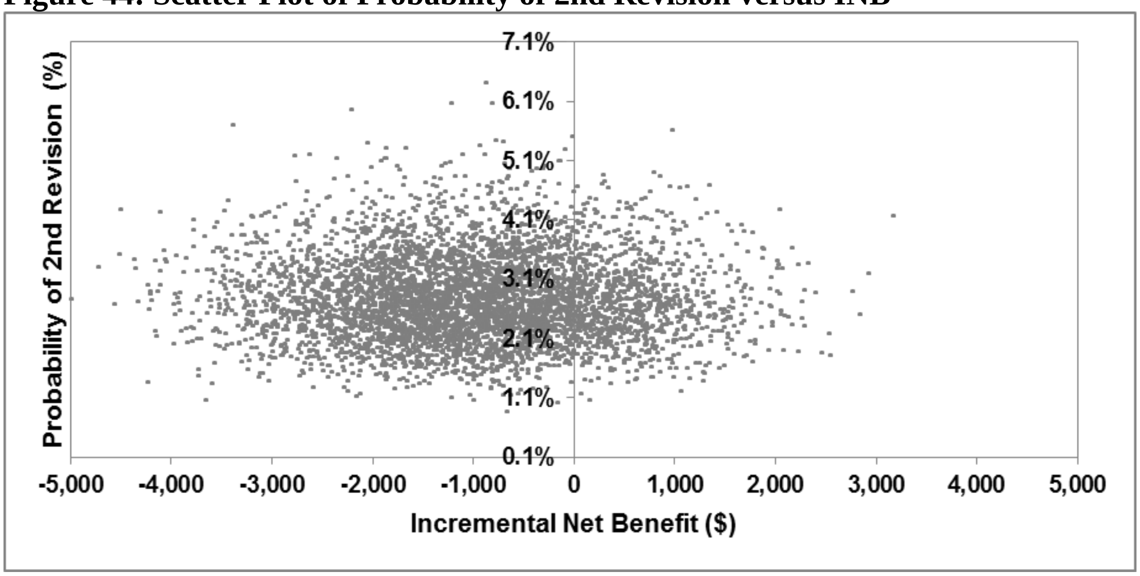


Figure 45: Scatter Plot of Probability of Successful Revision versus INB

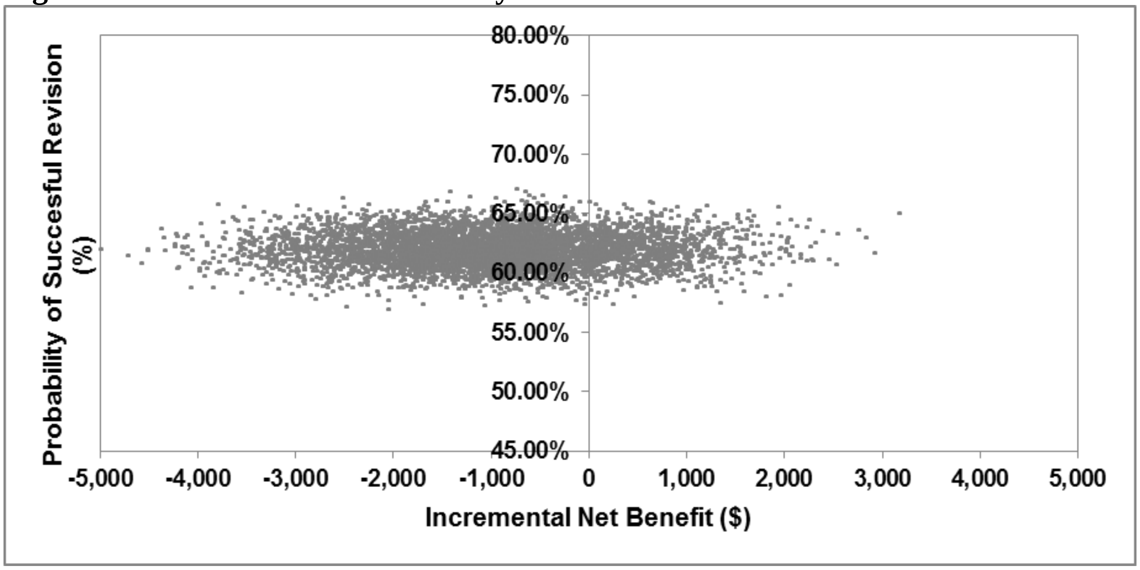


Figure 46: Scatter Plot of Probability of Failure or Complications versus INB