THE DEVELOPMENT AND APPLICATION OF A NORMATIVE FRAMEWORK FOR CONSIDERING UNCERTAINTY AND VARIABILITY IN ECONOMIC EVALUATION

A thesis submitted for the degree of Doctor of Philosophy

by

Douglas Coyle

Health Economics Research Group

Brunel University

December 2004

ABSTRACT

The focus of this thesis is in the development and application of a normative framework for handling both variability and uncertainty in making decisions using economic evaluation. The framework builds on the recent work which takes an intuitive Bayesian approach to handling uncertainty as well as adding a similar approach for the handling of variability.

The technique of stratified cost effectiveness analysis is introduced as an innovative, intuitive and theoretically sound basis for consideration of variability with respect to cost effectiveness. The technique requires the identification of patient strata where there are differences between strata but individual strata are relatively homogenous.

For handling uncertainty, the normative framework requires a twofold approach. First, the cost effectiveness of therapies within each patient stratum must be assessed using probabilistic analysis. Secondly, techniques for estimation of the expected value of perfect information should be applied to determine an efficient research plan for the disease of interest. For the latter, a new technique for estimating EVPI based on quadrature is described which is both accurate and allows simpler calculation of the expected value of sample information. In addition the unit normal loss integral method previously ignored as a method of estimating EVPPI is shown to be appropriate in specific circumstances,

The normative framework is applied to decisions relating to the public funding of the treatment of osteoporosis in the province of Ontario. The optimal limited use criteria would be to fund treatment with alendronate for women aged 75 years and over with

previous fracture and 77 years and over with no previous fracture. An efficient research plan would fund a randomised controlled trial comparing etidronate to no therapy with a sample size of 640. Certain other research studies are of lesser value.

Subsequent to the analysis contained in this thesis, the province of Ontario revised there limited use criteria to be broadly in line with the conclusions of this analysis, Thus, the application of the framework to this area demonstrates both its feasibility and acceptability.

The normative framework developed in this thesis provides an optimal solution for decision makers in terms of handling uncertainty and variability in economic evaluation. Further research refining methods for estimating information value and considering other forms of uncertainty within models will enhance the framework.

Contents

Index of Tables and Figures

Note: For all tables and figures, data reported are from original analysis for this thesis unless stated otherwise.

Acknowledgements

I must acknowledge the support and encouragement from many friends, colleagues and family members. Without your assistance over the past four years this thesis would never have finally materialised! In particular, I wish to acknowledge the following:

To my family for their constant support and love over what has not always been an easy four years. To my kids for their patience and their constant questions about what a thesis is. To my wife Sheila for her constant love, encouragement and support even when she was at her lowest ebb. Special thanks for checking chapters as they appeared.

To my colleagues in Ottawa and elsewhere for their words of encouragement and support. To my own students who have inspired me to get this thesis complete even if I often failed to heed the advice I gave them. Special thanks go to Kathryn O'Grady for her meticulous proof reading of the final draft.

To Jeremy Oakley for his assistance in statistical notation with respect to Chapter 4 and to the anonymous referees who have commented on the various papers which have formed the basis of certain sections within this thesis.

To all my friends at the Health Economics Research Group who have made me feel welcome during my frequent visits over the past four years and have provided many pertinent comments when I have presented the content of this thesis.

To the University of Ottawa and the Ottawa Health Research Institute for their support in terms of financial backing and time. Special thanks go to George Wells and Andreas Laupacis.

Finally, I wish to give the most praise to my supervisors, Martin Buxton and Bernie O'Brien for their constant help. Martin has helped me in so many ways over the course of this thesis as well as encouraging me to commence these studies in the first place. Bernie O'Brien has given me great support and friendship over my time in Canada and it is to his memory that this thesis is dedicated.

Publications Relating to Thesis

A number of papers have been published which cover material within this thesis. They are outlined below in terms of the specific Chapter to which they relate.

Chapter 2

Two papers where I have been sole author outline the case for basing decision making on the expected value for incremental net benefit and criticise the WHO recommendations for probabilistic analysis.

> Coyle D. (2003) Determining the optimal combinations of mutually exclusive interventions: a response to Hutubessy and colleagues. Health Econ, vol 12, pp 159-62.
>
> Coyle D. (2003) WHO's better not best: appropriate probabilistic uncertainty analysis. Int J Technol Assess Health Care, Vol 19, pp 540-5.

Chapter 3

The framework for stratified cost effectiveness analysis and the initial analysis using data from the Mark et al. was published in Health Economics. I was the principal author and I am grateful for comments from my supervisors acting as co-authors.

> Coyle D, Buxton MJ, O'Brien B. (2003) Stratified cost-effectiveness analysis: a framework for establishing efficient limited use criteria, Health Econ, vol. 12, pp. 421-7.

Chapter 4

The case study used to demonstrate the framework for the handling of uncertainty was published in 2003 and I was the principal author. Co authors provided clinical expertise and conducted data collection.

> Coyle D, Barbeau M, Guttman M, Baladi J-F (2003). The economic evaluation of pharmacotherapies for Parkinson's disease treatment. Parkinsonism Relat Disord, vol. 9, pp. 301-7.

The use of the quadrature method for estimating EVPPI was first suggested in a poster presentation at the 2001 Society for Medical Decision Making. I was the principal author. Co authors provided clinical expertise and conducted data collection.

> 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.

The review of alternate measures of parameter importance was published in the Journal of Clinical Epidemiology. The analysis in the published paper has been revised slightly for the thesis and the conclusions are similarly modified. I was the principal author for this paper and I am grateful for comments from my supervisors acting as co-authors.

Coyle D, Buxton MJ, O'Brien BJ. (2003) Measures of importance for economic analysis based on decision modeling. J Clin Epidemiol, vol. 56, pp.989-97.

Chapter 5

I was responsible for the development of the OMERACT guidelines for economic evaluations of treatments for osteoporosis and was principal author of the published report. My co author provided assistance in developing these guidelines.

Coyle D, Tosteson ANA. (2003) Towards a reference case for economic evaluations of osteoporosis treatments. J Rheumato, vol. 68, pp. 31-36.

Chapter 6

I am solely responsible for the development of the economic model for osteoporosis detailed in Chapter 6. The model was first developed in 2000 and previous versions of the model have been used in three previous publications. Co authors provided clinical expertise and research support.

Coyle D, Lee KM. (2002) Evidence based economic evaluation: how the use of different data sources can impact results. In: Donaldson C, Mugford M, Vale L. (eds.) Evidence Based Health Economics, London, BMJ Books.

Coyle D, Cranney A, Lee KM, Welch V, Tugwell P. (2001) Cost effectiveness of nasal calcitonin in postmenopausal women: use of Cochrane Collaboration methods for meta-analysis within economic evaluation. Pharmacoeconomics, vol. 19, pp. 565-75.

Waldegger L, Cranney A, Man-Son-Hing M, Coyle D. (2003) Cost-effectiveness of hip protectors in institutional dwelling elderly. Osteoporos Int, vol. 14, pp. 243-50.

The utility values used in the economic model were derived from an Ottawa based study. The study was the basis of a master's thesis for which I was the thesis supervisor.

Cranney A, Coyle D, Pham BA et al. (2001) The psychometric properties of patient preferences in osteoporosis. J Rheum, vol. 28, pp. 132-7.

Chapter 1.

Introduction

Given the scarcity of health care resources, it is necessary to demonstrate that new therapies provide value for money in comparison with other potential interventions. Economic evaluation provides an analytical framework for assessing the costs and benefits of interventions thus providing information to facilitate decisions relating to resource allocation.

In interpreting the results of economic evaluation based on decision analysis, decision makers must consider two related but diverse issues: variability in results between potential patients and uncertainty in results for particular patients. The focus of this thesis is in how analysis should be conducted to permit decision makers to optimally consider these issues.

Hutubessy and colleagues (2001) argue that "little or no attention is paid to the question of how decision makers should interpret the results where uncertainty levels interlap". (p473). However, several articles have addressed the interpretation of results under uncertainty (Claxton et al. 2001, Felli and Hazen 1999, Meltzer 2001). A number of these articles have highlighted the irrelevance of focusing on inference with respect to such decisions within a public health care system. Such articles have proposed Bayesian approaches to deal with such issues which highlight the need to consider both the optimal treatment choice and the value to be obtained from further information (Claxton et al. 2001, Felli and Hazen 1999).

In contrast, to the recent work in considering uncertainty within economic analysis based on decision analysis, there have been few careful considerations of the concept of variability. Most developments in this area have focussed on the handling of variability in economic evaluations based within randomised controlled clinical trials (Hoch et al. 2002).

In this thesis a normative framework is developed for handling both variability and uncertainty in making decisions using economic evaluation. A normative framework relates to what decision makers ought to do given the circumstances facing them. Within this thesis, it is the decision maker with responsibility for making reimbursement decisions over new technologies which is considered. Thus, the normative framework relates to how such decision makers should deal with uncertainty and variability. The framework allows for other decision makers to not act optimally. The framework builds on the recent work which takes an intuitive Bayesian approach to handling uncertainty as well as adding a similar approach for the handling of variability.

In Chapters 2, 3 and 4 the normative framework is developed. In Chapter 2, definitions of the concept of uncertainty and variability are provided. In addition, a brief review of previous guidance on how they can be handled is conducted. In Chapter 3 a framework is developed for handling variability in economic evaluation. In this chapter, the methods for conducting stratified cost effectiveness analysis are developed and illustrated through a case study. In Chapter 4, the framework for handling uncertainty is developed. The framework is consistent with an intuitively

Bayesian approach to the consideration of uncertainty. The optimal treatment choice is assumed to be the treatment with the highest net benefit with further focus on the value of information to be obtained from refining parameter estimates within the economic evaluation. Methods for eliciting estimates of the expected value of perfect information are reviewed as are other methods of determining parameter importance from other disciplines.

The rest of the thesis involves the application of the normative framework to treatment decisions relating to the management of osteoporosis in Canada. Chapter 5 provides a review of previous economic evaluations in this area as well as a review of previous guidance for the conduct of these evaluations. In Chapter 6, details of an economic model for the evaluating treatment choices in Canada for osteoporosis are provided. Chapter 7 reports the methods and results of a stratified cost effectiveness analysis for the management of osteoporosis. Analysis is conducted to assess the optimal treatment choice for osteoporotic women in Canada based on their age and fracture history. The analysis presented in this chapter was the basis for a recent revision to reimbursement criteria for osteoporotic women in the province of Ontario. In Chapter 8, a formal value of information analysis is conducted to determine an efficient research plan relating to osteoporosis management. All potential research studies are considered and those with potential information value are identified. Ultimately, the greatest information value was found to relate to the conduct of randomised trials of therapies.

Chapter 9 presents the conclusions of this thesis. The normative framework is reviewed and argued to be applicable. The innovative features contained within the

thesis are highlighted and future areas of research are identified. The application of the framework to the management of osteoporosis demonstrates that even for complex disease processes application of the framework is feasible and the information provided by analysis can improve the efficiency of both health care provision and research funding.

Chapter 2.

Variability and Uncertainty

2.1 INTRODUCTION

In this chapter, formal definitions of uncertainty and variability are provided. In addition current practice with respect to the handling of these concepts within economic analysis are provided.

2.2 VARIABILITY

2.2.1 Definition

In economic analysis, traditional sensitivity analysis often fails to distinguish between first order uncertainty (variability) and second order uncertainty (knowledge uncertainty) (Hoffman and Hammonds 1994). Knowledge uncertainty (henceforth uncertainty) relates to the lack of confidence in a parameter estimate due to lack of knowledge and is discussed further in Section 2.3.

Variability relates to the randomness in the population. Unlike uncertainty, variability cannot be reduced through further information. Rather, variability is best handled by stratifying the population into more homogeneous groups. Thus, variability is of importance with respect to policy decisions in that economic analysis can facilitate stratification of potential recipients of treatment into two distinct populations; those for which therapy is cost effective and those for which it is not.

Variability can occur within all input parameters within an economic analysis. For example, clinical effectiveness can vary by the number of prevalent risk factors (Klotzbuecher et al. 2000), costs may vary by place of residence (Wiktorowicz et al. 2001) and utilities may vary by age or cultural grouping (Cranney et al. 2001a, Coyle et al. 1999).

In statistical terms, variability relates to the standard deviation in the estimate of a single parameter which is the square root of the expectation of the square of the difference between an individual value and the expected value. As variability is inherent in a population it can-not be reduced. Rather, by splitting a sample into smaller groupings with similar characteristics the standard deviation within these groups will be smaller.

2.2.2 Previous Recommendations Related to Variability in Economic Analysis

2.2.2.1 Introduction

In this section, the extent to which there has been previous guidance on handling variability in economic analysis is ascertained[1]. Two forms of documents were reviewed: first, national guidelines for the conduct of economic analysis which may or may not have an explicit link to decision making; secondly, leading textbooks and other guidance statements for the conduct of economic analysis.

2.2.2.2 National Guidelines for the Conduct of Economic Analysis

A review of published literature identified nine national guidelines for the conduct of economic evaluation.

---

[1] The review is not comprehensive but is designed to provide insight to the extent to which issues of heterogeneity have been recognized as important.

Of the nine identified, four ignore the issue of heterogeneity[2]. In addition, the Australian guidelines for economic analyses submitted by the pharmaceutical industry recognise that the cost effectiveness of an intervention will vary by the indication for which it is used and that analysis should be conducted for all indications (Commonwealth Department of Human Services and Health 1995). However, there is no recognition that cost effectiveness can vary within a specific indication.

Two guidelines propose limiting consideration of heterogeneity in cost effectiveness to sub-groups for which there is a clinical rationale for any differences. National Institute of Clinical Excellence guidance to manufacturers making submissions to technology appraisals explicitly addresses the issue of what is an acceptable sub-group analysis (NICE 2001)[3]. NICE limits such analysis to sub-groups where there is "a sound biological a priori rationale" and only "where there is evidence that clinical effectiveness or cost-effectiveness may vary between such groups". The Canadian Coordinating Office of Health Technology Assessment (CCOHTA) guidelines for pharmacoeconomic analysis present the argument for conducting sub-group analysis within a traditional frequentist statistical framework (CCOHTA 1997). It is argued that sub-group analyses should only be presented if such sub-groups were identified within a clinical study protocol and then again only if the economic analysis by sub-group was "statistically sound".

[2] The four guidelines, which ignore heterogeneity were: guidelines for economic analysis proposed by the pharmaceutical industry in the USA (Clemens et al. 1995), guidelines funded by the Spanish Ministry of Health and Consumption (Rovira and Antonanzas 1995), Italian guidelines proposed by the Centre for Health Economics at the Mario Negri Institute (Garattini et al. 1995) and guidelines developed for use by the province of Ontario in Canada (Detsky 1993).

[3] Newer NICE (2004) guidelines developed since the publication of many papers relating to this thesis are discussed in Chapter 9.

Only, two guidelines explicitly recognise the need to conduct analysis for different sub populations. In the Dutch guidelines developed by the Health Insurance Council, it is recommend that analysis be conducted for populations which have "differences in effectiveness, costs and/or other parameters" (Ziekenfondsraad 1999). Similarly, Belgian guidelines proposed by the Belgian Society of Epidemiology (BSE) recommend that analysis be conducted for groups with "differential effectiveness, costs or preferences" (BSE 1995).

Thus, existing national guidelines differ in the extent that variability in cost effectiveness is considered. Certain guidelines explicitly state the need to conduct analysis for subpopulations. However, none of the available guidelines discuss the concurrent handling of uncertainty and variability. Nor, do the guidelines provide an explicit framework for handling variability.

2.2.2.3 Other Forms of Guidance

In the textbook by Drummond and colleagues, the issue of variability in cost effectiveness across the potential patient population is mentioned only briefly (Drummond et al. 1997). In discussing the critical assessment of published studies, it is noted that studies may present results across a range of patient characteristics. No further comment is given relating to the conduct and interpretation of such analysis.

Drummond and Jefferson (1996) have provided guidelines for the reporting of economic evaluations aimed at authors and peer reviewers. These have been adopted

by the British Medical Journal and by other peer reviewed clinical journals. The guidelines do not provide any advice with respect to heterogeneity in outcomes.

The report of the Panel on Cost Effectiveness in Health and Medicine convened by the US Public Health Service provides a more detailed discussion of handling variability in economic analysis (Gold et al. 1996). In the section relating to the target population of a cost effectiveness analysis, it is noted that there may be both "effectiveness" and "cost" sub-groups. "Effectiveness sub-groups" are identifiable groups for which the effectiveness of an intervention is likely to vary. Similarly, "cost sub-groups" are identifiable groups for which the cost of an intervention is likely to vary. It is noted that analysis by sub-groups may be of more use to decision makers though the lack of precision involved in such analyses should be considered. The role of modeling in the analysis of sub-groups is noted.

The report of the ISPOR Good Research Practices Task Force contains an explicit recommendation in relation to heterogeneity within a study population (Weinstein et al. 2003). The report states that errors can occur in the interpretation of the results of a study if heterogeneity is ignored. The recommendation of the report is that "when appropriate, modelled populations should be disaggregated according to strata that have different event probabilities, quality of life and costs."

2.3 UNCERTAINTY

2.3.1 Definition

Uncertainty relates to the lack of confidence in a parameter estimate due to lack of knowledge. Knowledge uncertainty can be addressed through better measurement;

although further knowledge may not lead to less uncertainty for either or both the input parameter and the outcome measure. Knowledge uncertainty unlike variability can be represented by probability distributions.

Uncertainty can occur with any parameter within a decision model. Uncertainty over the expected value of an input parameter will lead to uncertainty over the expected values of outcomes within an economic analysis: uncertainty propagation. Reduced uncertainty around input parameters will thus reduce the uncertainty around outcomes.

Uncertainty around the expected value of an input parameter can be expressed by the standard error of the population mean – which is equivalent to the standard deviation divided by the square root of the sample size. Thus, if the standard deviation remains constant over the sample of interest increasing sample size will reduce the standard error[4].

Thus, given the uncertainty around input parameters, decision makers must deal with uncertainty over the expected values of interest and hence the relative cost effectiveness of treatment alternatives.

[4] However, it should be noted that if the parameter of interest is truly heterogeneous then increasing the sample size will lead to false inference in that the expected value will be assumed appropriate for the whole population.

2.3.2 Previous Recommendations Related to Uncertainty in Economic Analysis

2.3.2.1 Sensitivity Analysis

Traditionally, in economic analysis, analysis of uncertainty has focussed on the use of deterministic sensitivity analyses[5]. In such analysis the value for one or more parameters is changed and the effect on the outcome of the interest (e.g. the incremental cost effectiveness ratio) is assessed. Deterministic sensitivity analysis can take a number of forms (e.g. simple analysis, analysis of extremes and threshold analysis) though the basic methods are consistent.

There have been three major criticisms of the use of deterministic sensitivity analysis.

First, the choice of variables and the subsequent range of values to apply deterministic sensitivity analysis appears subjective. There appears to be little theoretical or statistical basis to make such decisions and in a complex model the number of potential analyses may be excessively large.

The second major concern relates to how such analyses can be interpreted. The principal focus of analyses appear to be to assess the "robustness" of the base study result. However, robustness also has no objective meaning and decision makers are provided little guidance on the interpretation of such analyses.

The final concern with deterministic analyses is that they can provide a biased estimate of the outcomes of interest especially when the relationship between input

[5] A substantial number of reviews of the use of sensitivity analysis have been published and it was felt unnecessary to repeat such an exercise (e.g. Briggs et al. 1994, Briggs 2000)

parameters and outcomes are non linear (Thompson and Graham 1996). Deterministic sensitivity analysis ignores this issue.

Thus, given the concerns above there has been increasing interest in the use of probabilistic sensitivity analysis. Uncertainty around parameters can be represented by probability distributions. Probabilistic methods such as Monte Carlo simulation techniques use the probability distributions for parameters rather than point estimates to estimate the expected values of outcomes and their dispersion (Doubilet et al. 1985). Methods for assessing cost effectiveness within probabilistic analysis are detailed in Section 4.3.

2.3.2.2 Theoretical Frameworks for Handling Uncertainty

When making any investment decision, an individual generally has to consider the competing concepts of expected value and risk (O'Brien and Sculpher 2000). Individuals make decisions not solely on the expected returns but with allowance for the costs of risk bearing, a function of risk attitude, associated with each option. Thus, decisions will differ across individuals based on their different estimates of expected returns and differences in their risk attitude.

In traditional clinical decision making, decisions are based on classical statistical inference which takes an extreme position with respect to the trade off between uncertainty and expected value. Rather than ignoring the level of uncertainty and basing the decision on expected value, statistical inference places more weight on the level of risk associated with the decision than on the expected benefits to be obtained. Decisions are based on the ability to reject a null hypothesis. Inference is

often based on the adoption of a type 1 error rate of 5%: i.e. a requirement that the risk of falsely rejecting the null hypothesis is less than 5%.

Many have suggested that methods for considering uncertainty in clinical decision making can be applied to decisions based on economic analysis. In defining a null hypothesis, a test statistic is required. Incremental net benefit (INB) has been suggested as an appropriate test statistic for assessing statistical significance within economic evaluation (Stinnett and Mullahy 1998, Zethraeus et al. 2003a)[6]. Thus, a null hypothesis may relate to the INB of a new therapy compared to current practice being less than or equal to 0. If the expected value of the INB of the new treatment is positive with the probability of a negative INB being greater than 5% then the null hypothesis could not be rejected and the new treatment would not be considered optimal. A role for hypothesis testing within economic evaluation has been suggested by several authors (e.g. Huninck et al.1998, Briggs 2000, Zethraeus et al. 2003a)

Stochastic league tables, an alternate approach to considering uncertainty in decision making, has been developed by the World Health Organization (WHO) (Hutubessy et al. 2001, Baltussen et al. 2002). This approach requires the adoption of Monte Carlo simulation analysis to identify the probability that a certain program will be included in an optimal mix given the uncertainty around the program's expected costs and benefits. Those programs with the highest probability of inclusion should be funded. However, it has been demonstrated that this methodology can lead to potential inefficiencies arising through the dependence of such probabilities on

[6] Methods for estimating incremental net benefit are discussed in section 4.3.

decisions relating to other programs and the failure to consider the opportunity costs of obtaining increased health benefits. In the example cited by WHO, it is possible that the "optimal" mix of interventions neither maximizes incremental net benefit nor falls within the desired budget constraint (Coyle 2003a, 2003b)

Despite the preponderance of articles relating to either statistical inference or confidence intervals and cost effectiveness analysis, there are convincing arguments to ignore such concerns when making allocative decisions under uncertainty. Arrow and Lind (1970) in their seminal article provide three potential positions with respect to whether public investment decisions should be consistent with private decisions in their consideration of risk alongside expected returns.

The first position is that within a public investment decision risk should be considered as it is in private decisions and that the costs of risk bearing should be considered with respect to the trade off with expected return. The second position is that governments should ignore risk in decision making as governments make a great number of investment decisions and thus is able to pool risks across these decisions. The third position is that individual and societal preference for risk are not necessary consistent and governments should determine a national policy with respect to risk preference for public investment decisions.

The first and second positions have been shown to lead to the same decision making criteria – that is that public investment decisions should be made based solely on expected returns (Arrow and Lind 1970). This occurs due to the negligible costs of risk bearing with respect to public investment decisions.

Claxton and Posnett (1996) further developed these arguments specifically in relation to determination of optimal treatment choice. They argued that if the objective of the health care system is to maximize health gain then decisions over optimal treatment choice should be made solely on the basis of the expected value of the net benefits from treatment. The rationale behind this argument is that a choice must be made between the treatments available – hence the treatment which has the highest expected returns should be chosen. This is further justified in that the opportunity costs associated with basing the decision on expected values are symmetrical.

Acceptance of the relevance of expected value decision making does not preclude the need to consider uncertainty within economic analysis of health interventions. This may preclude any role for hypothesis testing in determining optimal treatments. However, consideration of uncertainty is important for determining which further information should be collected to facilitate a reconsideration of this decision at a further date (Claxton and Posnett 1996, Felli and Hazen 1998).

The position taken within this thesis is broadly in line with that of Claxton and has been argued to be consistent with a Bayesian approach to decision making (O'Hagan and Luce 2003). Uncertainty over input parameter values will be handled by conducting probabilistic analysis through the use of Monte Carlo simulation. The optimal treatment choice will be based on expected values with uncertainty over the incremental net benefit being considered in terms of the value of further information.

The position is normative rather than positive as it reflects what decision makers ought to consider rather than what they may actually consider. Thus, techniques for determining the optimal trade off between uncertainty and expected returns such as portfolio theory are not considered further (O'Brien and Sculpher 2000).

2.4 CONCLUSION

The focus of this chapter has been to both define uncertainty and variability in the context of economic evaluations of health technologies as well as highlight how they have been handled previously. Despite their major differences previous analyses and guidelines have tended to argue for both concepts to be considered similarly through the conduct of sensitivity analysis.

The normative framework developed in Chapters 3 and 4 recognises the inherent differences between uncertainty and variability. In Chapter 3 a framework for the explicit consideration of variability is developed whereby analysis is stratified by patient characteristics. In Chapter 7, this framework is applied to an economic evaluation of treatments for osteoporosis.

In Chapter 4 a framework for the explicit consideration of uncertainty is developed. Despite the argued irrelevance of statistical inference with respect to economic evaluation, methods for representing the degree of uncertainty within an economic evaluation have been developed. In Chapter 4, these are discussed and illustrated using a case study outlined in Section 4.2. In further sections of Chapter 4, methods for determining the value of further information are detailed and demonstrated. In

Chapter 8, methods for assessing the value of information within stratified analysis (as defined in Chapter 3 and applied in Chapter 7) are discussed.

Chapter 3.

Handling Variability in Economic Analysis

3.1 INTRODUCTION

In Chapter 2, variability was defined as the lack of homogeneity in outcomes amongst a target patient population. Thus, the answer to the question "Is this treatment cost-effective?" will usually be "It depends" because the economic value is conditional upon *who* receives *what* therapy and under *what* circumstances. Governments and other payers for health care have recognized that gaining a better understanding of the heterogeneity between eligible patients in terms of effectiveness and cost, provides a basis for restricting technologies to specific patients. Restrictions may be based on clinical evidence and possibly economic evidence.

For example, in the Canadian province of Ontario, the government can reimburse a new medicine onto the public formulary in the category of "limited use" where physicians are required (but not legally bound) to prescribe the drug only for patients who meet certain clinical or demographic criteria (Laupacis 2002)[7]. Therefore the intent of a Limited Use Criteria (LUC) policy is to restrict the public subsidy of a medicine to a sub-group of those patients for whom it is licensed with the aim of improved value for money. Similar reimbursement policies can be found in the UK with the National Institute for Clinical Excellence (Rawlins 2004) and in Australia with the Pharmaceutical Benefits Advisory Committee (Glasziou and Mitchell 1996).

---

[7] In Ontario, 20 out of 37 drugs newly listed in provincial formularies between March 1999 and February 2001 were subject to limited use criteria.

The focus of this chapter is the derivation of a normative framework for handling variability in economic analysis. As a normative framework, the focus will be on what decision makers ought to do given potential heterogeneity; rather than how decision makers currently consider this issue. Section 3.2 contains the major focus of the chapter; the development of a framework which can be used to define and quantify the efficiency gains from stratification inherent in Limited Use Criteria (LUC) based upon heterogeneity between patients in terms of costs, outcomes or both (Coyle et al. 2003a)[8]. In Section 3.3, the framework is applied to a previously published study to demonstrate the potential efficiency gains from stratification.

3.2 FRAMEWORK FOR STRATIFIED COST EFFECTIVENESS ANALYSIS

3.2.1 Introduction

In Chapter 2, a review of current guidance on the conduct of economic analysis found little consideration of appropriate methods for considering variability in economic analysis. The report of the ISPOR Good Research Practices Task Force contains the most explicit recommendation (Weinstein et al. 2003)[9]. However, the report does not contain a detailed discussion of appropriate methods for conducting such analyses[10].

---

[8] The framework I have developed for handling variability was originally presented to the UK Health Economics Study Group in 2001 and has subsequently been published in Health Economics (Coyle et al. 2003a). I am grateful to the constructive comments made on previous versions of the framework especially by my supervisors, by Dr John Brazier and by anonymous referees for the journal.

[9] Note, that the report of the task force post dates the original HESG paper where the framework outlined in this section was first introduced and the electronic pre publication of the framework.

[10] Published articles do often contain recognition of the need to consider variability. For example, Kuntz and Goldie (2002) and Zaric (2003) both comment that a major assumption within a Markov model is that the population under consideration is homogenous. However, a framework for handling heterogeneity has not been provided.

The framework detailed in this section uses the concept of net benefit (Stinnett and Mullahy 1998) to identify the optimal stratification of patients, and to allow the quantification of the potential gains from stratification. The framework developed in this chapter can be referred to as stratified cost effectiveness analysis (Coyle et al. 2003a).

The framework allows for consideration of circumstances whereby the net benefit gains may not be fully realized and allows calculation of the associated net benefit loss. The first circumstance is where elements of a proposed efficiency-based stratification for LUC (e.g. age, gender) may be contested on grounds of equity. The second circumstance is the practical problem of the extent to which health care providers will adhere to the LUC and only prescribe to those patients where net benefit is positive. This can be referred to as leakage.

In Section 3.3, an application of the framework is conducted using data from a published example to illustrate how gains in net benefit from stratification can be calculated. The application will illustrate how the proposed framework permits an estimation of net benefit loss associated both with the imposition of different equity constraints and alternate levels of leakage.

3.2.2 Defining the Framework

Consider the situation where there is a new treatment available (t_1) for a particular condition where the traditional treatment (t_2) is still available. Thus, we wish to determine for which groups of patients, the use of t_1 is cost effective.

E_{t_1} is defined as the expected value of health benefits (e.g. QALYs) from treatment t_1 and C_{t_1} as the expected value of costs. The net monetary benefit for t_1 is defined as:

$$NB_{t_1} = \lambda * E_{t_1} - C_{t_1}$$

where

λ = *a decision maker' s maximum willingness to pay for a unit of health benefit*

Incremental net benefit (INB) for t_1 when compared to t_2 is defined as:

$$INB_{t_1 t_2} = \lambda * (E_{t_1} - E_{t_2}) - (C_{t_1} - C_{t_2})$$

Now the cost and/or effectiveness of treatments can vary by particular patient characteristics (e.g. age, gender, risk). Consider the situation where cost and effectiveness can vary by two factors: j (j=1,2 ... J) and k (k=1,2... K). Patients can thus be described as belonging to one of J*K specific strata defined by their characteristics relating to j and k. Define $NB_{jk_{t_1}}$ as the net monetary benefit of treatment t_1 for the jk^{th} cohort which can be defined as:

$$NB_{jk_{t_1}} = \lambda * E_{jk_{t_1}} - C_{jk_{t_1}}$$

The incremental net benefit (INB) for t_1 when compared to t_2 for treating patients in the jk^{th} cohort is defined as :

$$INB_{jk_{t_1 t_2}} = \lambda * (E_{jk_{t_1}} - E_{jk_{t_2}}) - (C_{jk_{t_1}} - C_{jk_{t_2}})$$

By taking summation over the jk cohorts weighted by the potential number of patients in each cohort (n_{jk}) we can define Total Net Benefit (TNB) for t_1 when compared to t_2 as:

$$TNB_{t_1t_2} = \sum_j \sum_k INB_{jk_{t_1t_2}} * n_{jk}$$

The incremental net monetary benefit statistic is used in preference to incremental cost effectiveness ratios because it allows calculation of the monetary values of gains and losses from alternate limited use criteria (LUC). In the situation where TNB_{t_1} >0, without stratification the optimal treatment choice would be to treat patients with t_1. However, an efficient LUC would limit treatment with t_1 to those cohorts where the expected value of the incremental net monetary benefit was positive.

The total net benefit ($TNB_{s(jk)}$) when restricting therapy to those cohorts where INB>0 can be expressed as follows:

$$TNB_{s(jk)} = \sum_j \sum_k INB_{jk_{t_1t_2}} * n_{jk} \quad \forall_{jk} \quad where\ INB_{jk_{t_1t_2}} > 0$$

$$where\ s = stratification$$

Therefore, the net benefit gain from stratification (Δ_STNB) will be equivalent to the negative of the sum of the population weighted net benefit in the cohorts where net benefit is negative.

$$\Delta_S TNB = TNB_{s(jk)} - TNB_{t_1}$$
$$= -\sum_j \sum_k INB_{jk_{t_1t_2}} * n_{jk} \quad \forall_{jk}\ where\ INB_{jk_{t_1t_2}} < 0$$

3.2.3 Consideration of Efficiency/Equity Tradeoffs

In addition to a desire for efficiency in the provision of health care treatments, decision makers may have concerns over the access to treatment and the distribution

of health outcomes. Thus, decision makers may have to consider both efficiency and equity when considering an appropriate allocation of resources.

The framework outlined above permits the explicit consideration of a trade-off between equity and efficiency by determining the opportunity cost of an equity position. The more bases for stratification the greater the opportunity for efficiency gains[11]:

$$TNB_{s(jk)} \geq TNB_{s(j)} \geq TNB$$

Thus, if decision makers reject stratification based on a certain criterion for equity reasons, there will be an associated opportunity cost which can be expressed as a reduction in net benefit. The loss in total net benefit can be defined as Δ_ETNB: the opportunity cost of considering equity. Based on the above, where cost effectiveness is assumed to vary by two factors (j and k), the loss from not stratifying on the basis of k and only stratifying on the basis of j can be expressed as follows

$$\Delta_E TNB = TNB_{s(jk)} - TNB_{s(j)}$$

If a decision maker chooses not to stratify based on one patient characteristic, then Δ_ETNB is the decision maker's minimum willingness to pay for equal access to the therapy regardless of this characteristic: the opportunity cost of equity with respect to this characteristic.

[11] This of course assumes that the process of stratification is costless. Obviously the requirement to obtain data for stratification and the need to repeat analyses for all strata comes at increased costs and at some point such costs may outweigh the benefits from stratification.

3.2.4 Impact of Non-adherence

LUC are argued to be a "rigid" method for influencing the use of health technologies (Laupacis 2002). However, for this to be true, criteria must be adhered to. LUC though are rarely strictly enforced. Often a physician honour system exists whereby when prescribing therapy, physicians must state whether a patient meets the required criteria for reimbursement. This can lead to physician non adherence with LUC. This is referred to as leakage.

The framework above assumes adherence to the LUC. For LUC based on a strict efficiency criterion, leakage will necessarily reduce the net benefit from stratification[12]. If we define leakage (l_{jk}) as the proportion of patients who receive treatment in each cohort where net benefit is negative ($INB_{jk_{t_1t_2}}<0$), then the total net benefit from stratification given leakage ($TNB_{s(jk)}| L$) will be:

$$TNB_{s(jk)} \mid L = \sum_j \sum_k INB_{jk_{t_1t_2}} \quad \forall_{jk} \; where \; INB_{jk_{t_1t_2}} > 0$$
$$+ \sum_j \sum_k l_{jk} * INB_{jk_{t_1t_2}} \quad \forall_{jk} \; where \; INB_{jk_{t_1t_2}} < 0$$

Thus, the net benefit loss from leakage (Δ_LTNB) is:

$$\Delta_L TNB = TNB_{s(jk)} - TNB_{s(jk)}|L$$
$$= -\sum_j \sum_k l_{jk} * INB_{jk_{t_1t_2}}$$

[12] The previous sections follow the aims of the thesis by presenting a normative framework for decision makers. Thus, in the proposed framework decision makers faced with health care funding decisions are assumed to act rationally. However, it should be noted that other decision makers such as individual physicians and patients have interests which may conflict with those of the decision maker. The impact of behaviours associated with such interest should be considered within the framework. Such interests can lead to non adherence to limited use criteria as considered in this section.

Given the existence of leakage, there may be an alternative stratification basis ($s_l(jk)$) which will return a higher net benefit ($TNB_{s_l(jk)}|L$):

i.e. $TNB_{s(jk)}|L < TNB_{s_l(jk)}|L$

This can occur when the loss from leakage is substantively greater than the net benefit gain from a particular patient cohort.

In applying the estimate of leakage in practice, it would be necessary to specify a prior probability distribution for leakage proportions in each of the negative net benefit cohorts. Leakage can take different forms. Leakage can be indiscriminant; when physicians prescribe therapy regardless of the LUC; and discriminant; when the probability of leakage is greater for strata which are close to meeting the criteria and smaller for strata which are more distant. Discriminant leakage can be either accidental due to difficulties in assessing LUC or deliberate[13]. The form of leakage will influence the net benefit gained from stratification and will affect the likelihood that any revised stratification bases will be optimal.

3.3 EXAMPLE OF A STRATIFIED COST EFFECTIVENESS ANALYSIS

3.3.1 Case Study

In the previous section a framework for considering variability within economic analysis is outlined. To illustrate the impact of stratification policies, data are used from a published economic study of thrombolytic treatments for acute myocardial infarction (Mark et al. 1995). The study compared streptokinase which was the then

[13] Practitioners may be willing to prescribe therapy to patients who almost meet the criteria because they envision the benefits to these patients are greater than for those who are more distant. They may also be willing to prescribe to patients when they believe the potential penalties from failing to adhere may be minimal.

standard of care with tissue plasminogen activator (t-PA). The incremental cost per life year gained (ICER) of t-PA for all patients was $32 678 though this estimate varied by both patient age and location of infarction (anterior or inferior). Table 3.1 details the eight age and infarct location cohorts and their associated incremental costs and incremental life years for t-PA compared to streptokinase. For this illustrative example we assume a patient population with 100 patients in each age-location cohort due to the absence of data on the actual distribution of patients. In practice, the actual distribution of patients must be taken into account[14].

3.3.2 Stratification of the Potential Patient Population

Ultimately, each of the potential cohorts will be classified as belonging to one of two strata; those for whom a therapy is cost-effective and those for whom it is not. The cost effectiveness of t-PA for each cohort is expressed in terms of INB. This is derived as follows:

$$INB_{t\text{-}PA} = \sum_j \sum_k \left[\lambda * \left(IE_{jk_{t-PA}}\right) - \left(IC_{jk_{t-PA}}\right)\right] * n_{jk}$$

where

j = 1 for inferior infarct, 2 for anterior infarct

k = 1 for age <40, 2 for age 41-60, 3 for age 61-75, 4 for age >75

n_{jk} = number of patients in the jk^{th} cohort in the candidate population of users of this treatment

$IEjk_{t\text{-}PA}$ = incremental life-years gained per patient in the jk^{th} cohort when treated with t-PA rather than streptokinase

$ICjk_{t\text{-}PA}$ = incremental cost ($) per patient in the jk^{th} cohort, when treated with t-PA rather than streptokinase

[14] If the actual numbers are unknown then this must be treated as stochastic.

Table 3.1: Incremental Costs, Life Years and Net Benefit for t-Pa Compared to Streptokinase per Patient

Location-Age Cohort	Incremental Life-Years Gained	Incremental Cost ($)	Net Monetary Benefit($) per patient		
			(λ=$25 000)	(λ=$50 000)	(λ=$100 000)
Inferior <40	0.014	2845	-2495	-2145	-1444
Anterior <40	0.023	2845	-2270	-1694	-543
Inferior 41-60	0.038	2845	-1894	-944	958
Anterior 41-60	0.057	2845	-1419	7	2859
Inferior 61-75	0.102	2845	-293	2259	7362
Anterior 61-75	0.138	2845	608	4060	10965
Inferior >75	0.175	2845	1533	5911	14667
Anterior >75	0.212	2845	2459	7763	18371
Average	0.095	2845	-471	1902	6649

λ = alternative monetary values of life years gained.
Shaded cells relate to cohorts with negative net benefit where treatment with t-PA is not optimal.

Source: Costs and life years gained - Mark et al. (1995)
Incremental net benefit – original analysis

λ = threshold (maximum) monetary value of life-year, which is assumed constant over each cohort[15]

3.3.3 Benefits from Stratification

Analysis follows the framework outlined in Section 3.2.1. Table 3.1 presents the mean net monetary benefit of therapy for each cohort of 100 patients based on a monetary value of a life year of $25 000, $50 000 and $100 000. The optimal LUC will vary by the monetary value of a life year (λ). In this example, the proportion of patients for whom t-PA is optimal increases as the threshold increases (Figure 3.1). However, this is dependent on the location of each cohort on the cost effectiveness plane[16].

The optimal LUC assuming a threshold of $50 000 is to give therapy only to patients with an anterior infarction aged over 40 and patients with an inferior infarction aged over 60. The gain in total net benefit (Δ_STNB) from adopting this LUC is $478 237 (the population weighted sum of all cohorts with negative net benefits[17]) (Table 3.2). This is equivalent to the opportunity loss of a policy whereby all patients received this therapy.

Figure 3.2 depicts Δ_STNB over alternative thresholds for the monetary value of a unit of effect (λ). In this example, Δ_STNB peaks at the incremental cost per life-year

---

[15] One could allow λ to vary by cohort. For instance, society may have a higher preference for health benefits for the young than the old. Thus, λ may differ by the strata's age profile.

[16] Consider if there was a cohort where treatment is cost saving and less effective. In this instance, as λ increases, the cost effectiveness of treatment within this cohort will fall.

[17] Δ_STNB = – [(0.014 + 0.023+ 0.038)*$50 000 – ($2845 + $2845 + $2845)] * 100
= $478 237

Figure 3.1: Proportion of Patients for whom t-Pa is Optimal by Threshold Value

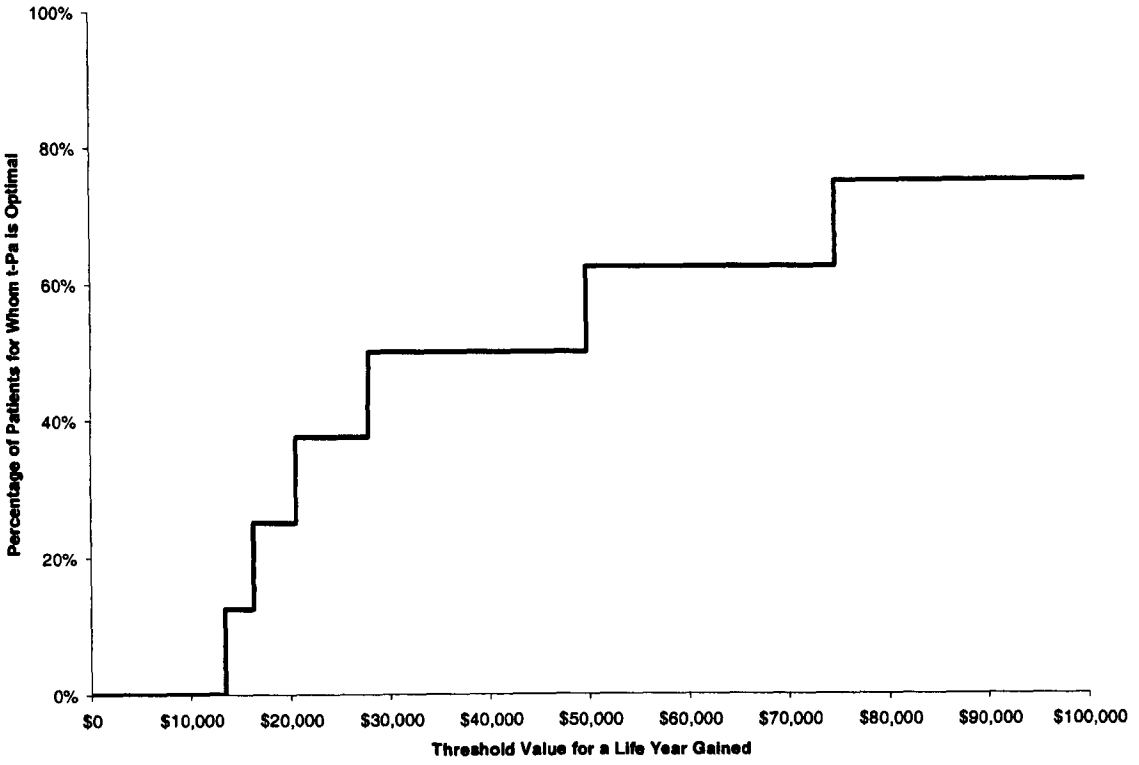


Figure 3.2: Net Benefit of Stratification by Threshold Value by Stratification Basis

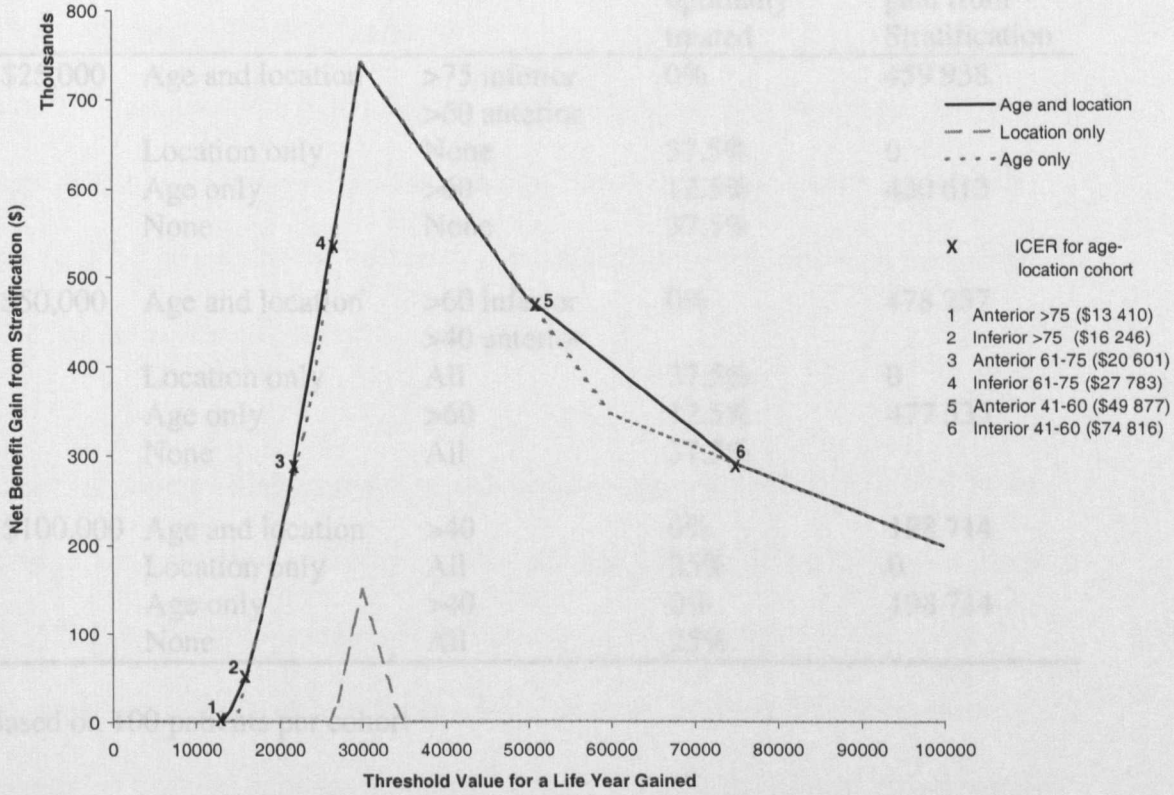


Table 3.2: Stratification of Patient Population Based on Maximizing Net Benefit

Threshold	Stratification Basis	Optimal Cohorts	Percent sub optimally treated	Net benefit gain from Stratification
$25,000	Age and location	>75 inferior >60 anterior	0%	459 938
	Location only	None	37.5%	0
	Age only	>60	12.5%	430 613
	None	None	37.5%	
$50,000	Age and location	>60 inferior >40 anterior	0%	478 237
	Location only	All	37.5%	0
	Age only	>60	12.5%	477 535
	None	All	37.5%	
$100,000	Age and location	>40	0%	198 714
	Location only	All	25%	0
	Age only	>40	0%	198 714
	None	All	25%	

Based on 100 patients per cohort

gained from t-PA over the whole patient population with kinks occurring at the incremental ratio for specific patient cohorts.

3.3.4 Efficiency/Equity Tradeoffs

A decision maker could potentially have equity concerns relating to stratifying on the basis of age or concerns relating to stratifying on the basis of disease (i.e. infarct location)[18]. Thus, it may be necessary to consider the opportunity cost from failure to stratify based on such concerns.

Table 3.2 presents the loss in net benefit of adopting restricted stratification bases incorporating only one of the risk factors assuming a value of $50 000 for a life year. Stratification based on location of infarct only would lead to a net benefit loss compared to the optimal stratification basis of $478 237 over the total patient population with an optimum policy of giving therapy to all patients. Stratification based on age only would lead to a net benefit loss compared to the optimal stratification basis of $702 over the total patient population with an optimum policy of giving therapy to all patients aged over 60. The loss in net benefit from restricted stratification varies significantly by the threshold value of a life year (Figure 3.3).

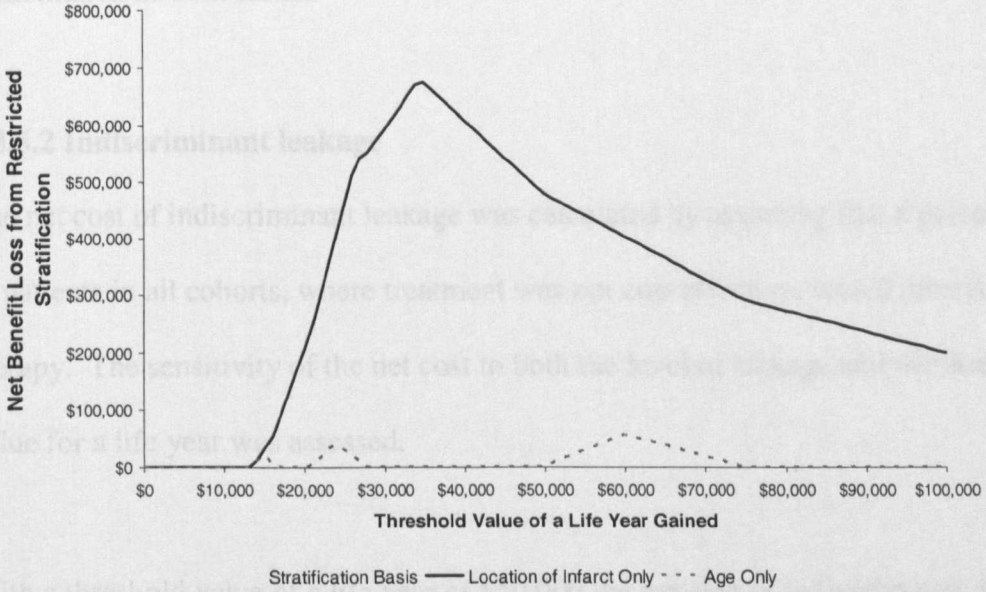
3.3.5 Impact of Non-adherence

3.3.5.1 Analysis

Two separate analyses were conducted to demonstrate the impact of leakage on the net benefit gain from stratification. First, analysis focused on indiscriminant leakage, by assuming leakage is equally likely to occur for all patients for whom therapy is

[18] Note, that the net benefit loss associated with failing to stratify by both age and infract location due to equity concerns is equal to $478 237, the net benefit from stratification.

Figure 3.3: Net Benefit Loss from Restricted Stratification Bases by Threshold Value



not cost effective. Secondly, analysis focused on discriminant leakage, by assuming leakage is restricted to "neighbouring cohorts"[19]. In reality, leakage is likely to be a combination of both forms.

3.3.5.2 Indiscriminant leakage

The net cost of indiscriminant leakage was calculated by assuming that a percentage of patients in all cohorts, where treatment was not cost effective, would receive therapy. The sensitivity of the net cost to both the level of leakage and the threshold value for a life year was assessed.

With a threshold value of a life year of $50 000 the net cost of indiscriminant leakage assuming a 10% level of leakage was $47 823[20]. The net cost of indiscriminant leakage is linear by the level of leakage and falls as the threshold value of a life year increases (Figure 3.4)[21]

Indiscriminant leakage can result in an optimal policy whereby therapy is denied to all patients. This can occur when the loss of net benefit from cohorts where treatment is not cost effective is greater than the net benefit in cohorts where treatment is cost effective. For example, if the threshold value of a life year was $20 000, the optimal stratification policy would be to allow t-Pa for patients aged over 70. The net benefit from such a stratification would be $206 000 (without

---

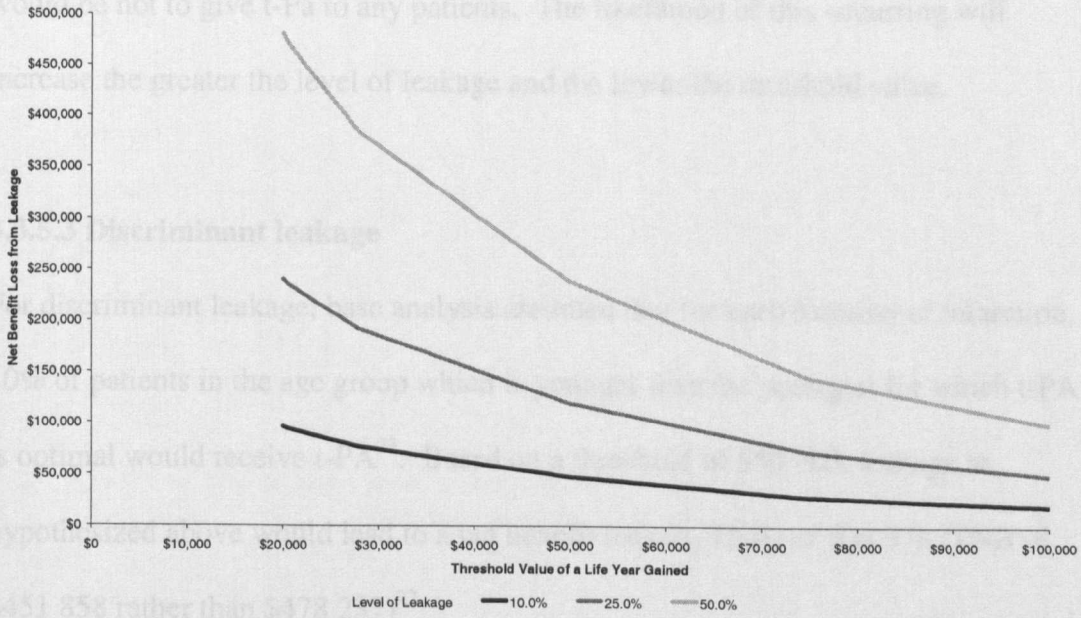
[19] That is, leakage is assumed to occur only for patients who just miss the cutpoint for therapy.

[20] $\Delta_L TNB = -0.1 * [\, INB_{\text{aged 41-60, inferior infarction}} + INB_{\text{aged } \leq 40\text{, inferior infarction}} + INB_{\text{aged } \leq 40\text{, anterior infarction}} \,] * 100$

$= -0.1 * [-944 - 2145 - 1694\,] * 100 \quad = \$47\,823$

[21] This is because for all cohorts therapy is more effective. Thus as λ increases the net benefit for each cohort increases. Hence, the consequences of leakage falls.

Figure 3.4: Net Benefit Loss from Indiscriminant Leakage by Level of Leakage and Threshold Value of a Life Year



stratification t-Pa would not be cost effective). However, if indiscriminant leakage was 25%, the net benefit loss from leakage would be $241 000. Thus, net benefit loss from leakage outweighs any gains from stratification and the optimal policy would be not to give t-Pa to any patients. The likelihood of this occurring will increase the greater the level of leakage and the lower the threshold value.

3.3.5.3 Discriminant leakage

For discriminant leakage, base analysis assumed that for each location of infarction, 10% of patients in the age group which is younger than the youngest for which t-PA is optimal would receive t-PA[22]. Based on a threshold of $50 000, leakage as hypothesized above would lead to a net benefit loss (Δ_LTNB) of $26 378 (TNB of $451 858 rather than $478 237)[23].

With discriminant leakage, Δ_LTNB varies by the cost effectiveness threshold as well as by the hypothesized level of leakage (Figure 3.5). Δ_LTNB is linear with respect to the level of leakage. However, Δ_LTNB does not have a monotonic relationship with λ. As λ increases, Δ_LTNB peaks at the ICER for individual cohorts. This occurs because at values just below an ICER for a cohort the net benefit loss from individuals in that cohort receiving therapy will be minimal, as therapy is almost cost effective. For values just above the ICER, net benefit will rise as the cohort to which therapy may leak will be comparatively less cost effective to treat.

---

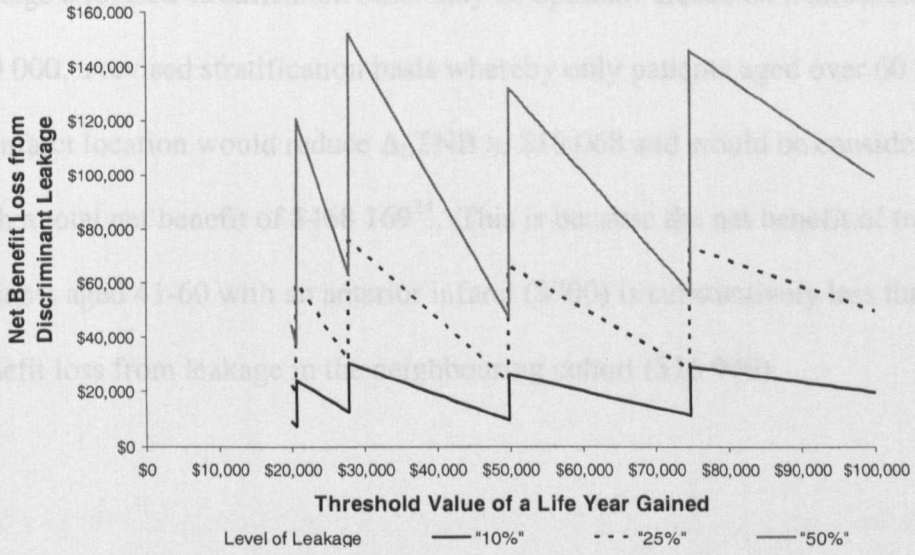
[22] For example, with a threshold of $50 000, it is cost-effective to restrict therapy to patients with an inferior infarction who are aged over 60. Thus under the assumption given, 10% of patients with an inferior infarction who are aged 41-60 are assumed to receive t-Pa.

[23] $\Delta_L TNB = -0.1 * [(INB_{\text{aged 41-60, inferior infarction}} + INB_{\text{aged } \leq 40, \text{ anterior infarction}}] * 100$

$= -0.1 * [-944 - 1696] * 100$

$= \$26\,378$

Figure 3.5: Net Benefit Loss from Discriminant Leakage by Level of Leakage and Threshold Value of a Life Year



The case study demonstrates the situation where in the presence of discriminant leakage a revised stratification basis may be optimal. Based on a threshold value of $50 000, a revised stratification basis whereby only patients aged over 60 regardless of infarct location would reduce Δ_LTNB to $10 068 and would be considered optimal with a total net benefit of $468 169[24]. This is because the net benefit of treating patients aged 41-60 with an anterior infarct ($700) is substantively less than the net benefit loss from leakage in the neighbouring cohort ($16 940).

3.4 CONCLUSIONS

Variability within economic analysis has tended to be ignored by authors of textbooks or guidelines relating to economic analysis. Consequently, few economic analyses in health care currently include a systematic consideration of the effects of heterogeneity between patients in terms of costs, effects or both.

In this chapter, a normative framework for considering variability within economic analysis is described. The framework allows full consideration of the variability between groups. Based on the principle of net benefit maximization, and allowing decision makers to consider alternative money values for units of effectiveness, the framework allows identification of the optimal criteria for restricting the use of technologies by identifying those groups for which therapy is of positive net benefit.

---

[24] $\Delta_L TNB = [-0.1 * INB_{\text{aged 41-60, inferior infarction}} + 0.9 * INB_{\text{aged 41-60, anterior infarction}}] * 100$
$= [-0.1 * -944 + 0.9 * 7] * 100$
$= \$10\,068$

Optimal cohorts for therapy could be determined based on the incremental cost per QALY of therapy compared to usual care for each cohort. However, an important advantage of the framework is that it permits quantification of the efficiency gains obtained from stratification as well as quantifying the opportunity costs of non-adherence and consideration of any equity efficiency trade offs. Although many studies report results in the form of incremental ratios, the net benefit statistic can easily be estimated assuming that incremental costs and outcomes are reported in a disaggregated form.

Defining strata for reimbursement is one approach to heterogeneity in cost-effectiveness data, but it should be recognized that stratification of study data on costs and effects has a risk of misclassification as sample size within cells is reduced. Further sampling can reduce uncertainty about parameter values within strata and differences between strata. The value of such further information can be assessed by estimating expected values using probabilistic analysis and applying methods described in detail in the following chapter. However, the decision to acquire more information may be considered as independent from decisions relating to optimization based on current knowledge.

The concept of leakage is introduced and explored within the analysis of the case study where assumptions were made concerning the level and type of leakage. In reality, the extent of all forms of non-adherence are likely to be unknown and to some degree random. Thus, leakage could be represented by a random variable which may represent decision makers' prior beliefs. Results can be updated as

further information is made available with the potential for revisions to the limited use criteria.

In this chapter, there is no discussion concerning what are and are not suitable grounds for the stratification of patients. In theory, if there were no costs associated with stratification and enforcing LUC then it would be optimal to incorporate all criteria by which the cost effectiveness of treatment varies. However, stratification will not be costless and the feasibility of applying and enforcing LUCs based on multiple characteristics may be limited.

Concerns for equity may limit the acceptance of LUC based purely on concerns for efficiency although this should be addressed by considering the trade off between efficiency gains and such concerns. However, there may be criteria for which there are no equity objections with respect to stratification but inclusion of these within a limited use policy may be either too problematic or too costly.

Others have raised ethical objections to the use of cost effectiveness analysis as the sole basis for making allocative decisions relating to health care (Department of Health 2001, Hadorn 1992). Methods for weighting health benefits have been suggested as a basis for allowing for equity concerns with respect to resource allocation (Bleichrodt 1997, Dolan 1998, Williams 1997). However, the methods proposed require the ability to determine an appropriate weight for equity concerns within a decision maker's utility function. The framework detailed in this chapter allows decision makers to focus on a more explicit value judgment: is the opportunity cost of equitable access justified?

In conclusion, the framework detailed in this chapter provides an intuitive solution to the issue of restricted access to therapy, which will allow maximization of efficiency gains as well as considerations of both equity and non-adherence. The framework is normative and permits maximization of net benefit as an objective function with or without equity concerns acting as a constraint.

Chapter 4.

Handling Uncertainty in Economic Evaluations

4.1 INTRODUCTION

In Chapter 2, the concepts of uncertainty and variability were defined. In Chapter 3, a normative framework for considering variability in economic analysis was proposed and illustrated by application to a published economic analysis. In this chapter, I will address methods for handling uncertainty.

The chapter focuses on outlining the theory and logic behind analysing uncertainty within economic analysis illustrated by a simple case study. The proposed framework is based on two fundamental questions that face decision makers when making decisions under uncertainty:

- Given knowledge uncertainty what is the optimal treatment choice for each patient group?
- Given uncertainty propagation into outcomes of interest, what further information should be collected to reduce the potential opportunity costs of uncertainty[25].

With respect to the former question, in Chapter 2 it was argued that decisions on funding health care interventions should be based solely on the expected value of the

---

[25] The opportunity cost of making a decision under uncertainty can be defined as the product of the probability that a decision maker chooses one therapy when another therapy is superior (Type 3 error) and the net benefit gain that would have been derived if the superior treatment was chosen. This is equivalent to the expected value of perfect information which is discussed in detail in Section 4.5.

net benefits of individual treatment options. However, despite the argued irrelevance of statistical inference with respect to economic evaluation as discussed in Chapter 2, methods for representing the degree of uncertainty within an economic evaluation have been developed. In Sections 4.3 and 4.4, these are described and illustrated using a case study outlined in Section 4.2.

The major focus of this chapter relates to the latter question, concerning further information requirements. In Section 4.5, alternate methods for estimating the value of perfect information concerning parameters are described. These methods can be complex and potential simpler methods adopted in other disciplines are discussed in Section 4.6. In Section 4.7, methods for determining the value of further information are detailed and demonstrated.

4.2 CASE STUDY

In this section, details of the case study used to illustrate methods for handling uncertainty in economic evaluation based on decision models are detailed. The case study is an evaluation of entacapone for the treatment of advanced Parkinson's disease (Coyle et al. 2003b).

Parkinson's disease is a chronic, progressive neurodegenerative disorder (Aminoff 1998). Usual care for Parkinson's disease is based on levodopa (L-dopa), the precursor of dopamine (Rivest et al. 1999). Approximately 50% of Parkinson patients experience reductions of motor control after five years of levodopa therapy (Lang and Lozano 1998).

Increasing the dose of levodopa to overcome this often leads to adverse events in the form of dyskinesias, hallucinations and confusion (Lang and Lozano 1998). Newer agents have been developed that further preserve levodopa in the periphery by inhibiting catechol-O-methyl transferase (COMT) metabolism (Rivest et al. 1999). Entacapone is a reversible inhibitor of COMT recently licensed for the treatment of patients with advanced Parkinson's disease (Brooks et al. 2000). Thus, the analysis focused on the cost effectiveness of entacapone as an adjunct to treatment with levodopa.

The analysis was performed through a Markov model developed within an Excel spreadsheet with Crystal Ball enhancement to allow Monte Carlo Simulation (MCS) (Crystal Ball 2000, Doubilet et al. 1985). The analysis was based on 5000 replications: i.e. 5000 estimates of the costs and QALYs associated with each treatment option were obtained by randomly sampling from each uncertain parameter's probability distribution.

Severity of Parkinson's disease was measured by the proportion of "off" time, a measure of the existence of motor fluctuations: the greater the off time, the less severe symptoms. The Markov model assumed three distinct states; mild/moderate disease (>25% "off" time per day), severe disease ($\leq$ 25%"off" time per day) and death. A cut-off point of 25% "off" time was chosen as this is associated with both a significant decrease in utility and a substantial increase in costs (Palmer et al. 2000). All patients were assumed to be in the severe state at onset of treatment (Figure 4.1). For patients receiving entacapone, transition probabilities were required for improvement from severe to mild/moderate disease, progression from mild/moderate

to severe disease and death. For patients receiving usual care only progression from severe disease to death was required

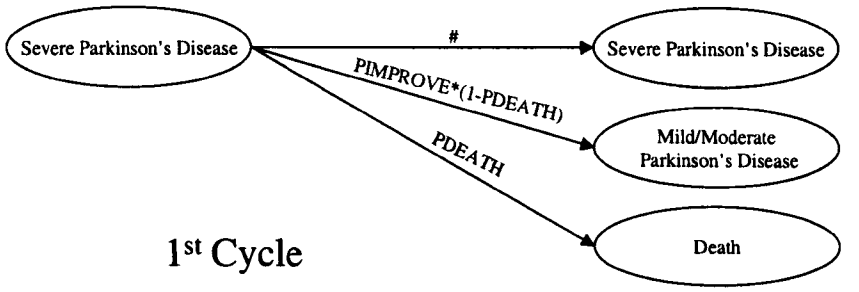
Analysis compared usual practice with and without the inclusion of entacapone. Usual therapy was assumed to include levodopa used in combination with other anti Parkinsonian medication. The model was based on a 6 month cycle. A five-year time horizon was chosen, which is relevant for a chronic disease like Parkinson's disease. All outcomes were discounted at 5%. Analysis was taken from the perspective of the health care system.

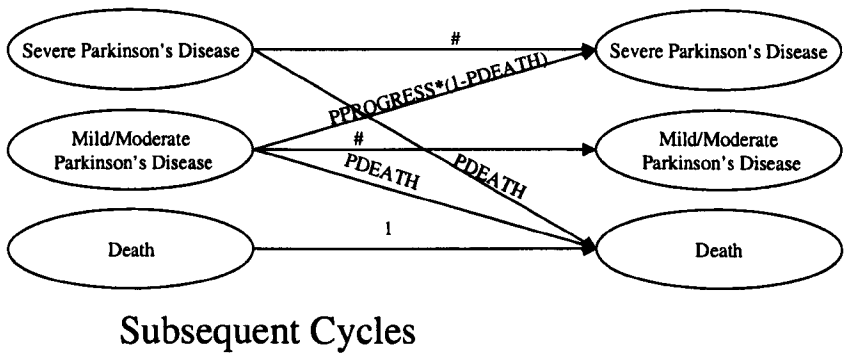
Input parameters and their associated probability distributions are detailed in Table 4.1. Drug costs were assumed fixed. The probability of mortality during each cycle was obtained from national population data and was also assumed fixed. This was felt justified as death rates were based on a large population size with little standard error and as they were based on population data the ability to obtain further information would be limited.

4.3 DEFINING COST EFFECTIVENESS WITHIN PROBABILISTIC ANALYSIS

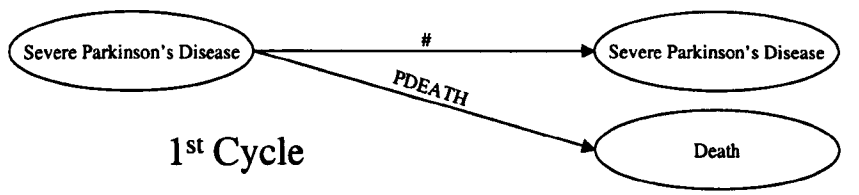
In this section, the algebraic formatting used to describe methods for handling uncertainty is detailed. This is followed by definitions of relevant outcomes for assessing cost effectiveness within probabilistic analysis.

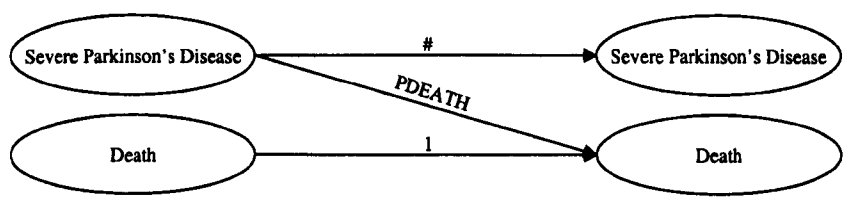
Figure 4.1: Design of Markov Model for Evaluation of Entacapone





a. Entacapone





b. Usual therapy

Note: Variable definitions are provide in Table 4.1

Table 4.1: Probability Density Functions for Input Parameters

Parameter		Mean	Probability Density Function
Probabilities			
Improvement from severe disease to mild disease with therapy	PIMPROVE	0.324	Beta (61, 27)
Progression from mild disease to severe disease	PPROGRESS	0.183	Beta (11, 49)
Probability of mortality	PMORT	0.032	Fixed
Utilities			
Mild disease	UMILD	0.75	Normal (0.75, 0.03)
Severe disease	USEVERE	0.64	Normal (0.64, 0.03)
Costs – mild disease			
Consultations	CCONSM	949	Normal (949, 189.25)
Hospital care	CHOSPM	1148	Normal (1148, 287)
Additional health care	CADDM	283	Normal (283, 70.75)
Costs – severe disease			
Consultations	CCONSS	2934	Normal (2934, 733.5)
Hospital care	CHOSPS	2567	Normal (2567, 641.75)
Additional health care	CADDS	578	Normal (578, 144.5)
Drug Costs			
Usual Care	CDRUGU	546	Fixed
Inclusion of entacapone	CDRUGE	1313	Fixed

Let $X = \{x_1, \ldots, x_d\}$ be the set of uncertain input variables, and T $(t_1, \ldots t_e)$ be the set of treatment options. An individual parameter within X is depicted by x_i. Within X, parameters can be grouped into parameter sub-sets; a sub-group of parameters is depicted by X_i. In both instances X^c denotes the complement set of input parameters: i.e. all members of X other than x_i or X_i[26].

E_{t_1} is the expected value of health benefits (e.g. QALYs) from treatment t_1 based on X and C_{t_1} is the expected value of costs. Expectations are based on data obtained from the Monte Carlo simulation (Doubilet et al, 1985).

Consider the case of two treatments t_1 and t_2. The cost effectiveness of a treatment (t_1) compared to an alternate treatment (t_2) can be presented in terms of the incremental cost per unit of health benefit (ICER). 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 (Stinnett and Paltiel 1997).

$$ICER_{t_1 t_2} = \frac{C_{t_1} - C_{t_2}}{E_{t_1} - E_{t_2}}$$

In the evaluation of entacapone, the incremental health care costs of entacapone were $1 200 with an incremental QALY gain of 0.07 (Table 4.2). This leads to an incremental cost per QALY gained (ICER) of $17 300.

The net benefit approach to depicting the cost effectiveness of a treatment option has been argued to have distinct advantages over the traditional ICER (Zethraeus et al.

[26] Consider the case where there are four parameters (x_1, x_2, x_3, x_4). A sub-group of parameters X_i may be defined as comprising x_1 and x_2. X^c the complement of X_i will comprise x_3 and x_4.

Table 4.2: Base Results of Economic Evaluation of Entacapone in Comparison with Usual Therapy

	Entacapone	Usual Therapy	Incremental Analysis
Costs	$52 900	$51 700	$1 200
QALYs	2.57	2.50	0.07
ICER			$17 300
Net Benefit (λ = $50 000)	$75 800	$73 500	$2 300

2003a, Stinnett and Mullahy 1998). Net benefit can be expressed in terms of net health benefits or net monetary benefits (Claxton and Posnett 1996). The net monetary benefit (henceforth NB) of a treatment defined in monetary terms is simply the difference between the expected effects and expect costs with effects weighted by the value of a unit of health benefit (λ). Thus, the expected net monetary benefit for an individual treatment is defined as:

$$NB_{t_1} = \lambda * E_{t_1} - C_{t_1}$$

Incremental net benefit (INB) can be calculated as the difference in net benefit between two treatment choices.

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

t* is defined as the optimal treatment choice. By definition, INB will be positive for the comparison between t* and any other potential treatment option.

For the base analysis, the INB of entacapone assuming a threshold value of a QALY of $50 000 was $2 300.

It is important to note that λ has two distinct interpretations. First, it can be interpreted as the shadow price of a unit of health benefit given a constrained budget (B). Thus, if we assume that the objective function of the decision maker is to maximise health benefit, the baseline decision is to choose the treatment with the highest expected net benefit (t*).

$$NB_{t*} = max_t \; E_X \left(NB_{t_1}\right)$$

Secondly, λ can be interpreted as the value society places on a unit of health benefit. Within a constrained health care budget, it may not be feasible to fund all interventions with a positive INB. Hence, the optimal decision will be to choose the treatment with the highest expected net benefit (t*) within the budget constraint.

$$NB_{t*} = max_t \; E_X \left(NB_{t_1} \right) | C_t < B$$

For this chapter, the first interpretation of λ is adopted. This is consistent with the interpretation of many authors who espouse Bayesian methods for handling uncertainty (e.g. Claxton and Posnett 1996, O'Hagan and Luce 2003, Felli and Hazen 1999). In Chapter 6, the impact of adopting the latter interpretation of λ is assessed.

The INB statistic allows the use of standard methods for determining statistical inference (Zethraeus et al. 2003a). However, the relevance of statistical inference for optimal decision making was questioned in Chapter 2. A further and potentially more important advantage is that NB allows adoption of a Bayesian approach to CEA in that the uncertainty around the cost effectiveness of treatment can be used to assess the value of collecting further information (O'Hagan and Luce 2003, Claxton and Posnett 1996).

4.4 REPRESENTING UNCERTAINTY IN ECONOMIC EVALUATIONS

4.4.1 Introduction

The normative framework adopted in this thesis requires that for decisions to be optimal in terms of maximising health benefits, decision makers should be concerned solely with the expected value of outcomes. However, several authors have

suggested methods for representing uncertainty concerning what is the optimal course of action. These methods are detailed in this section and their potential role in assisting decision makers discussed.

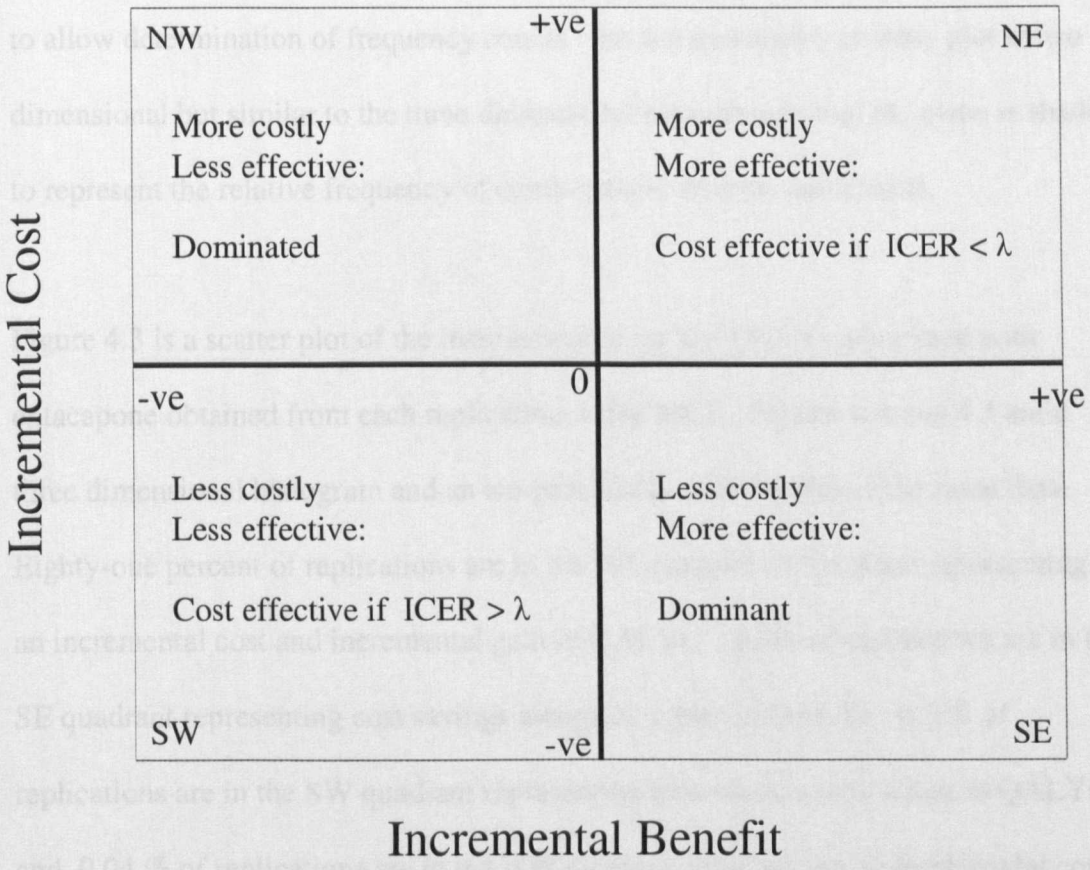
4.4.2 The Cost Effectiveness Plane and Scatter Plots

Figure 4.2 depicts the cost effectiveness plane (Black 1990, Briggs and Fenn 1998). 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 right hand or east half of the plane is associated with an effective treatment, whilst the upper or north half is associated with a more costly treatment option. 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) (Figure 4.2).

A scatter plot which involves plotting the incremental cost and incremental effect from each replication associated with each treatment option against one reference treatment (usually the least effective treatment) is a simple two dimensional means of graphically presenting the uncertainty over the results of a CEA. (Hunink et al. 1998). 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. Hence scatter plots are an empirical representation of the joint distribution of incremental costs and benefits.

Alternative graphic presentations have been suggested that are more complex visually (Hunink et al. 1998). A three dimensional histogram presents incremental

Figure 4.2: Cost Effectiveness Plane



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. An iso-probability contour plot is two dimensional but similar to the three dimensional histogram in that the plane is shaded to represent the relative frequency of combinations of costs and effects.

Figure 4.3 is a scatter plot of the incremental costs and QALYs associated with entacapone obtained from each replication of the MCS. Figures 4.4 and 4.5 are a three dimensional histogram and an iso-probability contour plot of the same data. Eighty-one percent of replications are in the NE quadrant of the plane representing an incremental cost and incremental gain in QALYs; 18.4% of replications are in the SE quadrant representing cost savings alongside a gain in QALYs; 0.5 % of replications are in the SW quadrant representing cost savings with a loss in QALYs; and, 0.04 % of replications are in the NW quadrant representing an incremental cost with an associated loss of QALYs.

The benefits from graphically presenting the result of a MCS are unclear. A scatter plot presents the distribution of costs and effects and allows a visual representation with respect to decision uncertainty. However, given the arguments in favour of basing decisions on expected values, there are unlikely to be any additional benefits from such presentations.

4.4.3 Credible Intervals

In recent years, there have been a large number of articles concerned with methods for estimating confidence intervals for ICERs based on data from randomized

Figure 4.3: Scatter Plot of Incremental Costs and QALYs with Treatment with Entacapone

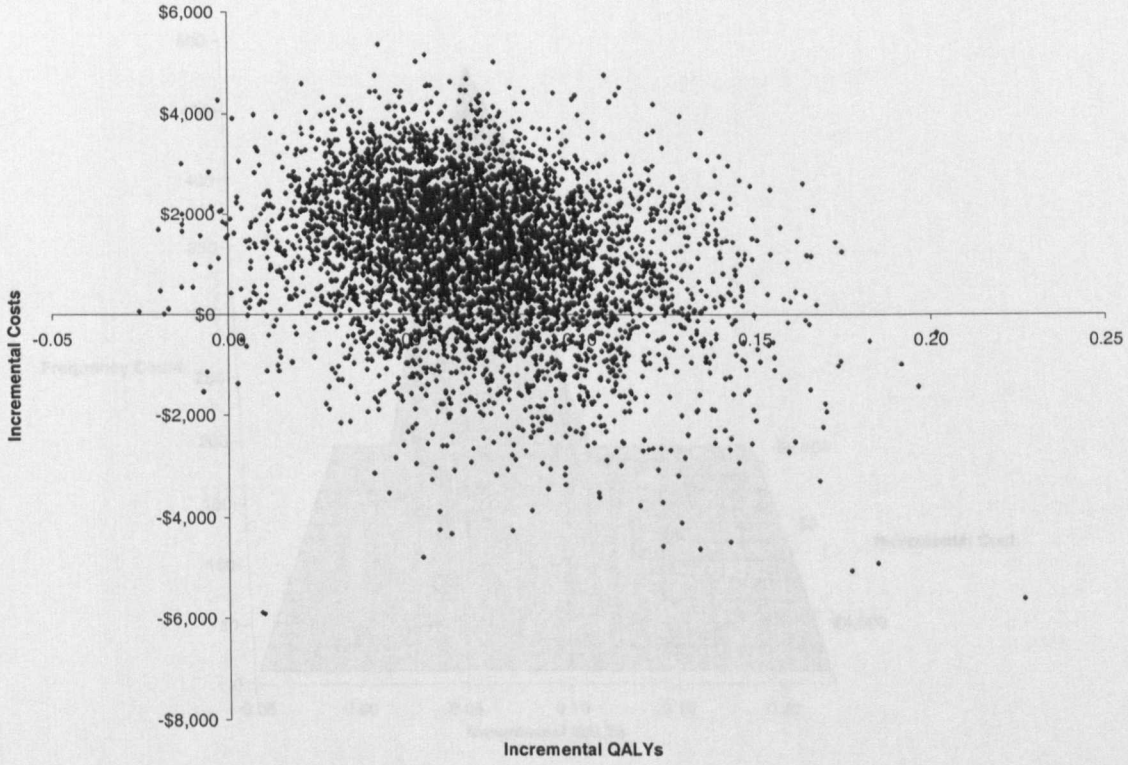


Figure 4.4: Three Dimensional Histogram of Incremental Costs and QALYs with Treatment with Entacapone

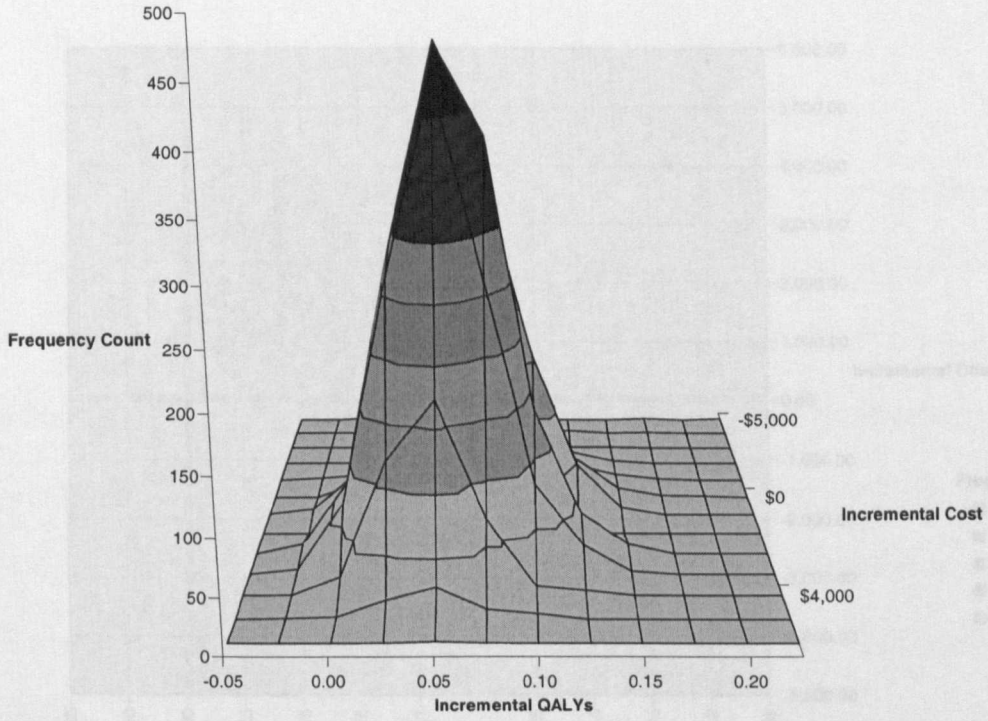
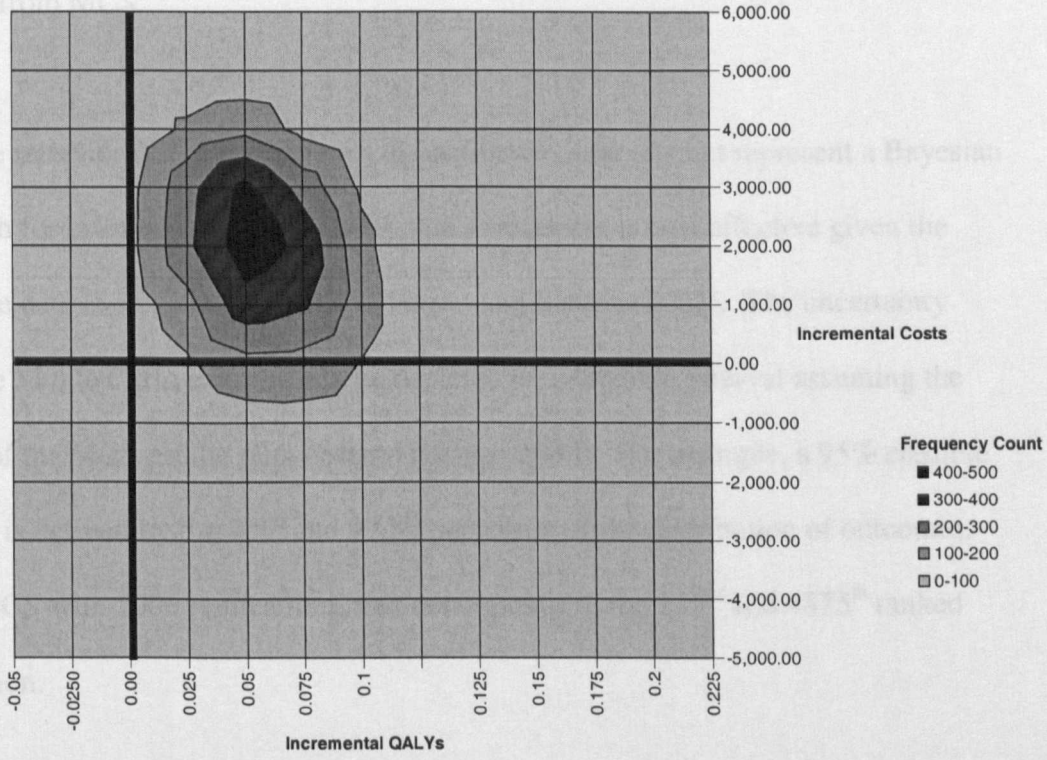


Figure 4.5: Iso Probability Contour Plot of Incremental Costs and QALYs with Treatment with Entacapone



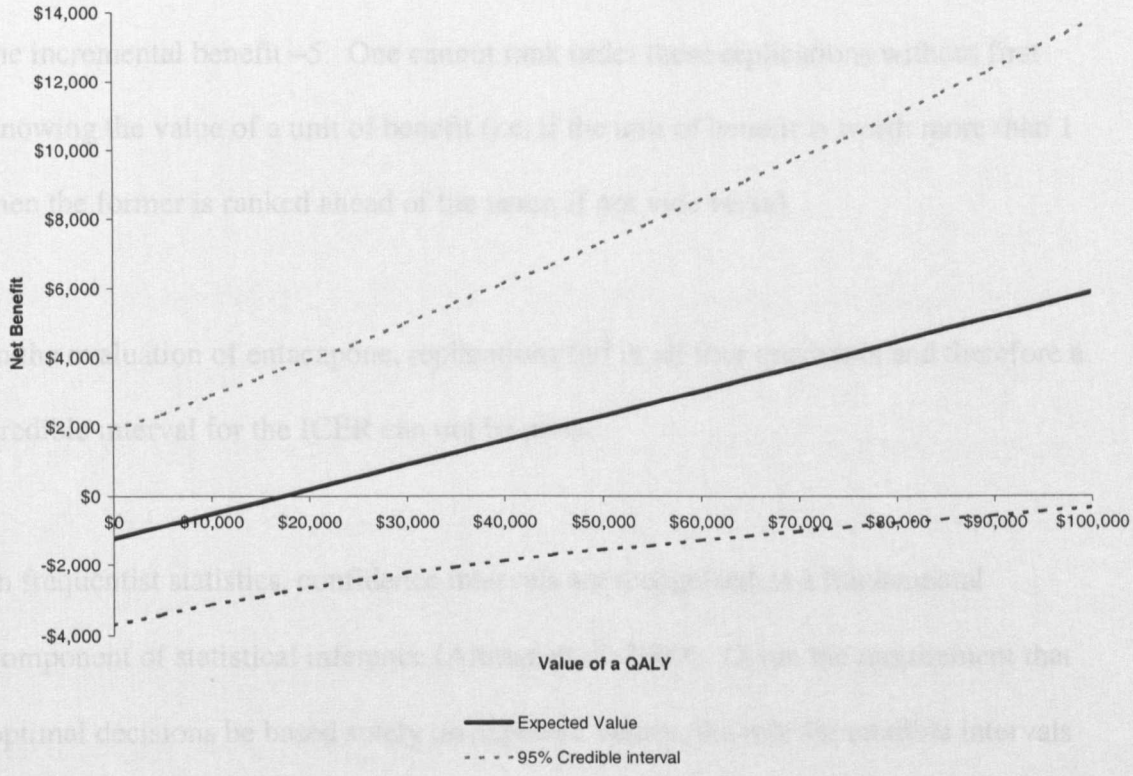
controlled trials (e.g. Briggs et al. 1999, Chaudhary and Stearns 1996, Willan and O'Brien 1996). These methods have mainly focused on a frequentist approach to statistical inference. In addition, the methods are unsuitable for the analysis of data derived from MCS.

Credible intervals (CI) are analogous to confidence intervals but represent a Bayesian approach for estimating the probability that a treatment is cost effective given the available data (Spiegelhalter 2000, O'Hagan and Stevens 2002). The uncertainty from the Monte Carlo exercise can be depicted by a credible interval assuming the output of the MCS can be rank ordered (Briggs 2000). For example, a 95% credible interval is defined by the 2.5^{th} and 97.5^{th} percentile of the distribution of outcomes. For a MCS with 5000 replications, this corresponds to the 125^{th} and 4875^{th} ranked replication.

The derivation of the 95% credible interval for the INB of a treatment is straightforward and will vary by λ. Figure 4.6 depicts the expected value for the INB of entacapone and the associated 95% credible interval for values of λ between \$0 and \$100 000. Note, that the expected value of net benefits is of course linear in λ but the credible interval is not. For λ equal to \$50 000 the INB of entacapone is \$2 300 (95% CI, -\$1 600 to \$7 300). Hence, based on the traditional rules of statistical inference, the null hypothesis could not be rejected and entacapone would not be considered optimal.

A credible interval for an ICER can only be estimated if the results of the MCS do not fall in both the SW and NE quadrant. This is because the ICERs from a MCS can

Figure 4.6: Incremental Net Benefit of Entacapone Compared with Usual Care



not be ranked if they fall in both quadrants. For illustration consider the following scenario. Assume two replications from a MCS, one where the incremental cost is $5 and the incremental benefit is 5 and another where the incremental cost is -$5 and the incremental benefit –5. One cannot rank order these replications without first knowing the value of a unit of benefit (i.e. if the unit of benefit is worth more than 1 then the former is ranked ahead of the latter, if not vice versa).

In the evaluation of entacapone, replications fall in all four quadrants and therefore a credible interval for the ICER can not be given.

In frequentist statistics, confidence intervals are recognised as a fundamental component of statistical inference (Altman et al. 2000). Given the requirement that optimal decisions be based solely on expected values, the role for credible intervals in economic evaluation is unclear and may have a very limited role as a summary statistic for depicting the level of uncertainty.

4.4.4 Cost Effectiveness Acceptability Curves

Cost effectiveness acceptability curves (CEAC) have been argued to provide a more intuitive way of presenting the uncertainty around the cost effectiveness of treatments than traditional confidence or credible intervals (van Hout et al 1994, Briggs and Fenn 1998). 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 (t_d) is optimal (i.e. has the greatest NB of all treatment options (T)) given a range of values for λ) (Fenwick et al. 2001, Briggs 2000, Lothgren and Zethraeus 2000, O'Hagan and Stevens 2002). The curve is presented

with an x-axis representing values of λ and a y-axis representing the proportion of replications from the MCS where NB is positive. The height of a CEAC at each value of λ, can be interpreted as the probability that the treatment is cost effective given the available evidence (Fenwick et al. 2001, O'Hagan and Stevens 2002).

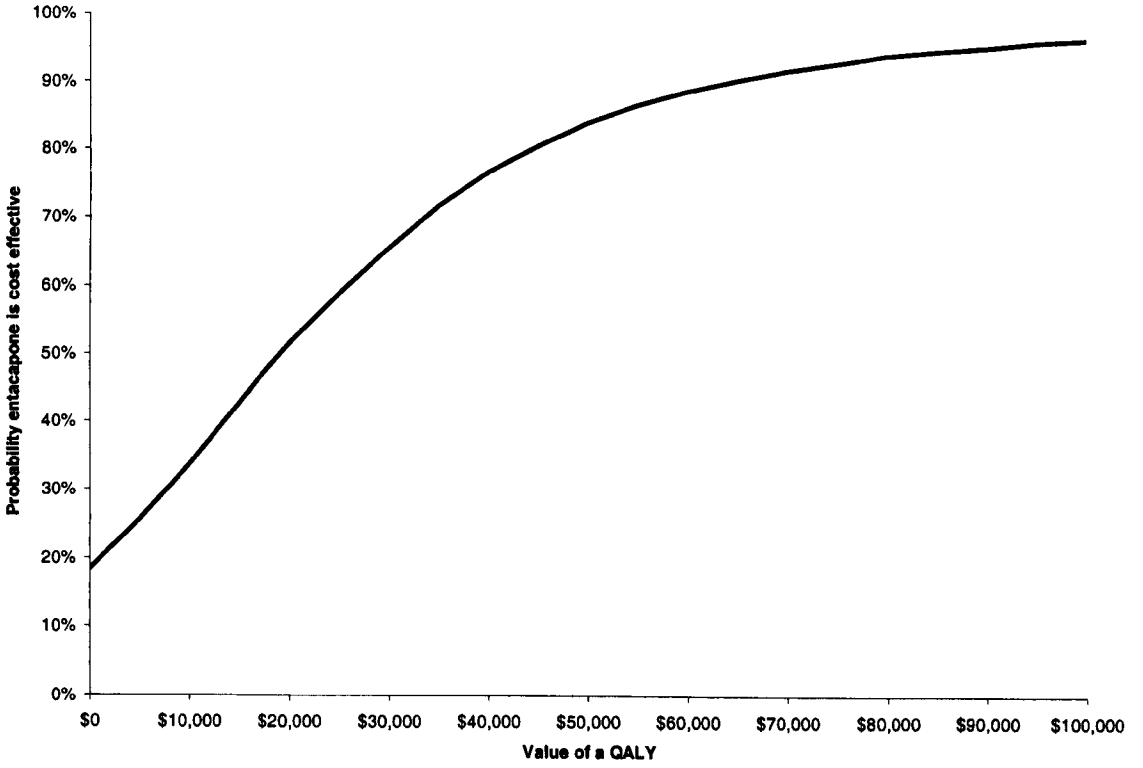
$$CEAC_{t_j}(\lambda) = p\left(NB(t_j, X) = max_T\{NB(T, X\}\right)$$

In a two treatment model, a CEAC will report the percentage of replications where one treatment is optimal, the complement of this percentage is the percentage of replications where the other treatment is optimal. In a multiple treatment model, the CEAC curve for each treatment can be represented on one graph (the sum of the height of all curves at each value of λ will be 1). In this instance, the y axis represents the proportion of replications where each treatment is associated the maximum NB (Fenwick et al. 2001).

Figure 4.7 depicts the CEAC for entacapone based on values of λ between $0 and $100 000.

A CEAC need not rise monotonically as λ increases. The shape of the CEAC is determined by the dispersion of the results of the MCS across the cost effectiveness plane. Replications in the NW quadrant do not contribute to the height of the curve as, regardless of the value of λ, therapy would never be cost effective given it is both costly and ineffective. In the SE quadrant, therapy is both effective and cost saving Thus, the contribution of the replications in the SE quadrant is constant regardless of the value of λ. The contribution of replications in the NE quadrant increases as λ

Figure 4.7: CEAC for Entacapone in Comparison with Usual Care



increases, whilst the contribution of replications in the SW quadrant decreases as λ decreases.

Thus, the boundaries of a CEAC are determined by the proportion of replications within the SE and NW quadrants; and the shape of the CEAC is determined by the relative contributions of the SW and NE quadrant. A curve, which monotonically increases as λ increases, occurs only when either there are no replications in the SW quadrant or the contribution of the NE quadrant increases at a greater rate than the contribution of the SW quadrant declines for all relevant threshold values. The latter is the case with treatment with entacapone as the curve does rise monotonically for values of λ less than \$100 000. For more extreme values the curve falls.

There are several problems with the use and interpretation of a CEAC particularly when more than two treatment options are being considered. If we accept that decision makers should make decisions solely on the expected values associated with each treatment option then the practical role of a CEAC appears negligible. The skewness of the distribution of net benefit may mean that therapies that are optimal based on expected values will not necessarily be the optimal treatment choice for the majority of replications. For the evaluation of entacapone, if λ was equal to \$18 000, entacapone should be considered cost effective with an expected net benefit of \$52. However, at this value of λ, entacapone is cost effective in only 48% of replications.

Problems with the interpretation of a CEAC are exacerbated when multiple treatment options are considered. This is discussed further in Appendix A in relation to the treatment of osteoporosis

An alternate but related approach to the CEAC is that of the cost effectiveness acceptability frontier (CEF) (Fenwick et al. 2001). The CEF presents, for each value of λ, the proportion of replications which show the optimal treatment choice as optimal.

$$CEF(\lambda) = p(NB_{t*} = max_T \{NB(T, X\})$$

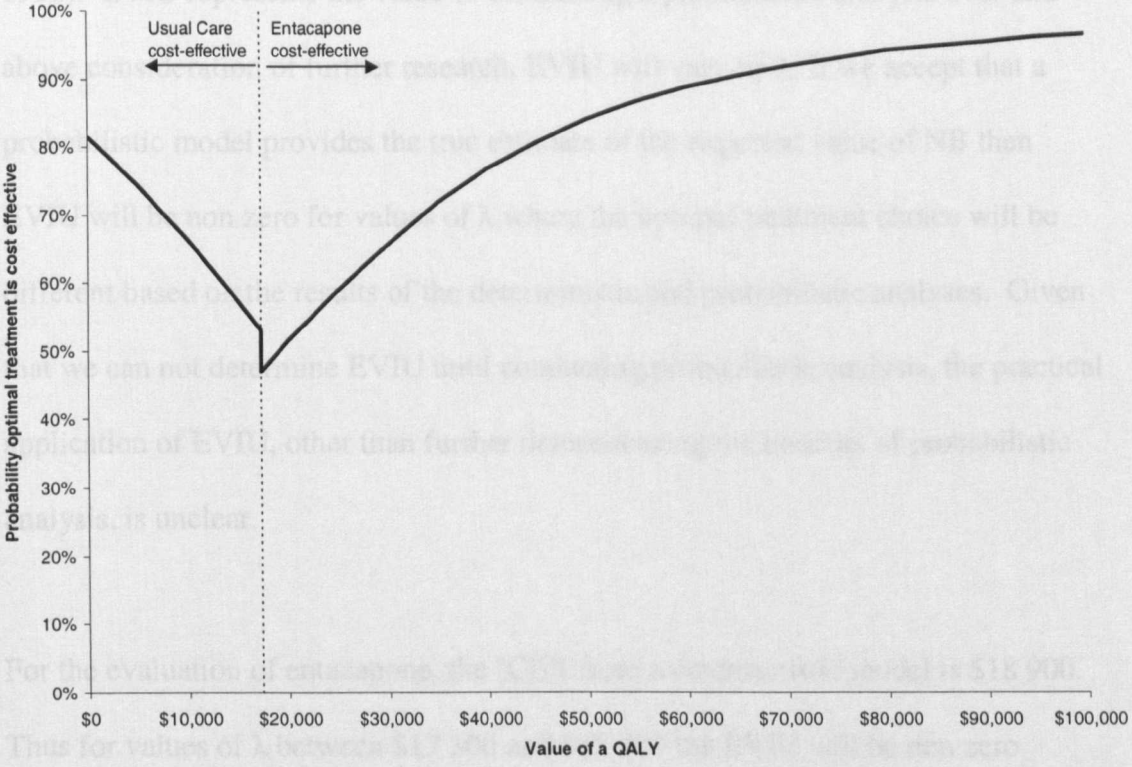
Figure 4.8 depicts the CEF for the treatment choice between entacapone and usual care. For values of λ below the ICER of $17 300, the CEF represents the probability usual care is cost effective and for values above the ICER, the CEF represents the probability that entacapone is cost effective. Note when λ equals $17 300, the CEF features a vertical line, a result of the lack of symmetry in the distribution of net benefit. This further illustrates the skewness of the distribution of net benefit and highlights how a CEF or a CEAC cannot be used to determine an optimal treatment choice.

4.5 ESTIMATING THE EXPECTED VALUE OF PERFECT INFORMATION

4.5.1 Introduction

Value of information analysis provides a framework for analysing uncertainty within economic analysis, by focussing on the value of reducing uncertainty through further information (Dakins et al. 1994). Such analysis adopts a Bayesian approach to sensitivity analysis (Felli and Hazen 1998, 1999). Within value of information analysis, there are two specific concepts to consider, the expected value of including uncertainty (EVIU) and the expected value of perfect information (EVPI).

Figure 4.8: Cost Effectiveness Frontier for Evaluation of Entacapone and Usual Care



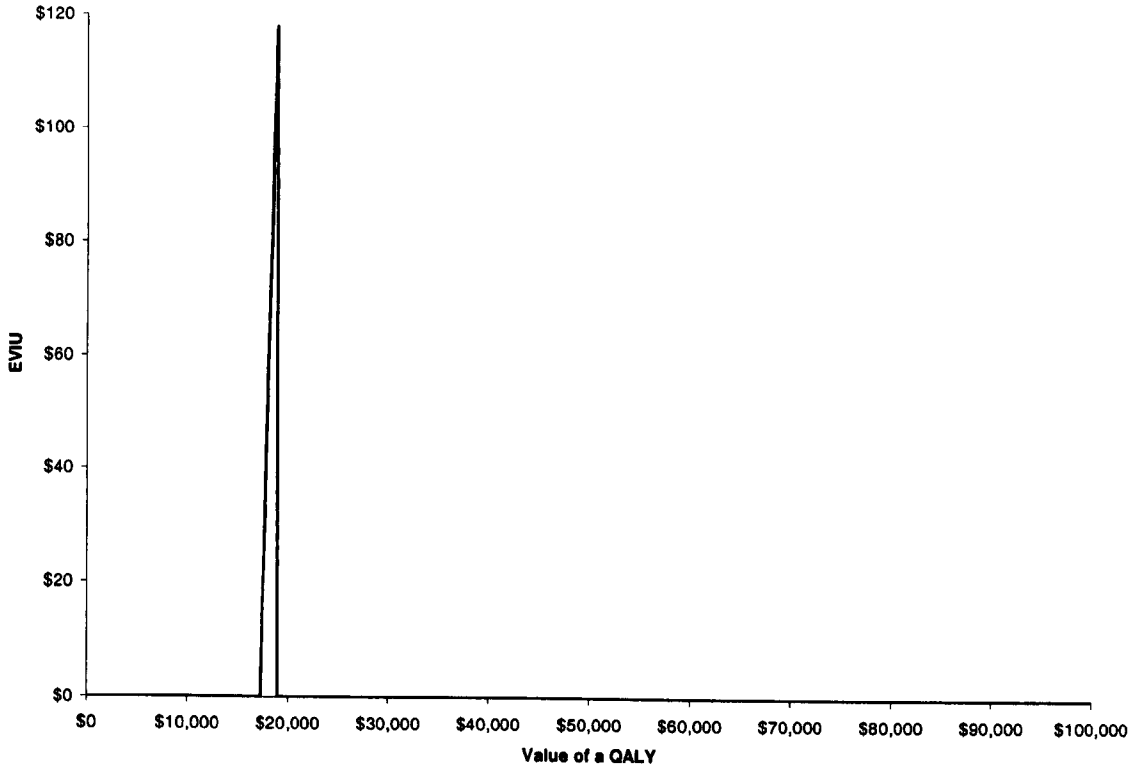
EVIU measures the difference in net benefit if the optimal treatment choice is based on a deterministic model rather than a probabilistic model (Morgan and Henrion 1990). EVIU represents the value of conducting a probabilistic analysis over and above consideration of further research. EVIU will vary by λ. If we accept that a probabilistic model provides the true estimate of the expected value of NB then EVIU will be non zero for values of λ where the optimal treatment choice will be different based on the results of the deterministic and probabilistic analyses. Given that we can not determine EVIU until conducting probabilistic analysis, the practical application of EVIU, other than further demonstrating the benefits of probabilistic analysis, is unclear.

For the evaluation of entacapone, the ICER from a deterministic model is $18 900. Thus for values of λ between $17 300 and $18 900 the EVIU will be non zero (Figure 4.9). For all other values of λ EVIU is zero.

Much of the focus of value of information analysis published in the health economics literature has been on the estimation of EVPI. EVPI is a 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 (Claxton and Posnett 1996, Felli and Hazen 1998, 1999). EVPI can be expressed as the 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.

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

Figure 4.9: EVIU for Analysis of Entacapone



Based on the results of a MCS, EVPI can be expressed as follows:

$$EVPI = E_X [max_t (NB_t)] - NB_{t*}$$

EVPI can be calculated for all parameters within a model (global EVPI). Alternatively, EVPI can be calculated for a partial set of input parameters (x_i or X_i.). This is termed the expected value of partial perfect information (EVPPI). Parameters for which the decision over optimal treatment is sensitive will have higher EVPPI, although for all parameters EVPPI will vary substantially by λ.

In this section, alternate methods for estimating EVPPI are described and applied to input parameters from the case study detailed in Section 4.2. Three of these methods have been described in the literature and applied to economic models (Claxton et al. 2001, Brennan et al. 2002a, Felli and Hazen 1998, 1999). One method has been used as a measure of global EVPI but has not previously been used as a measure of EVPPI (Claxton and Posnett 1996). The final method is comparable to one of the previously used methods but is argued here to be computationally more efficient (Coyle et al. 2003b).

EVPI and EVPPI are the most a decision maker should be willing to pay to alleviate uncertainty about input parameter(s)[27]. To estimate the optimal sampling frame for each individual and sub-set of parameters requires estimation of the expected value of sample information (EVSI) and the cost of sampling. This is discussed in detail in Section 4.7.

[27] This requires the simplifying assumption that the parameter(s) of interest is not pertinent to other treatment choices.

4.5.2 Mathematical Derivation of EVPPI

EVPPI for an individual parameter x_i is defined as:

$$EVPPI_{x_i} = E_{x_i} [max_t E_X | x_i (NB_t | x_i)] - NB_{t*}$$

EVPPI for a sub-group of parameters X_i is defined as:

$$EVPPI_{X_i} = E_{X_i} [max_t E_X | X_i (NB_t | X_i)] - NB_{t*}$$

EVPPI cannot be solved in a closed form. Thus, all methods of estimating EVPPI require integration using either Monte Carlo simulation or quadrature[28]. In the following sections five different proposed methods of estimating EVPPI are outlined.

The first two methods described are appropriate only in specific circumstances relating to the characteristics of the probability density functions of input parameters and their relationship with INB. In many instances the requirements required for these methods are not met and hence the methods are inappropriate for calculating EVPPIs for all input parameters. This is especially the case for Markov models[29]. Hence, it is necessary to adopt more complex methods, which can be applied in the general case. Three such methods are described. Two of these methods are based on the mathematical definition of EVPI and involve solving double integrals, neither of which are in closed form. The inner or nested integration involves estimating the incremental net benefit with different fixed values of X_i. The outer integration then determines EVPPI through integration across the probability density functions for X_i. An alternate method has been suggested which involves avoidance of the second

[28] Quadrature (or numerical quadrature) refers to numerical methods for estimating the area under the curve for functions which cannot be solved through integration. Common methods of numerical quadrature are the trapezoidal rule and Simpson's rule.

[29] In a Markov model, interactions of variables are compounded. For example, a transition probability may be used for multiple cycles. Thus, the relationship between outcomes and the transition probability will be non linear.

integral by assuming x_i is constrained to its expected value. This method is not based on the mathematical definition of EVPPI.

4.5.3 Methods of Estimating EVPPI[30]

4.5.3.1 Unit Normal Loss Integral Method

EVPI can alternately be described as the integral of the loss function (the probability density function for INB where INB<0) multiplied by the loss itself.

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

If the uncertainty around INB is normally distributed then there is a simple mathematical formula for the derivation of global EVPI (Chilcott et al. 2003a, Claxton and Posnett 1996). This approach is outlined as follows:

1. Derive a standardized distance D as the expected value of INB (μ_{INB}) less the break even point (i.e. INB = 0) divided by the standard deviation of INB (σ_{INB}).

$$D = \frac{\mu_{INB} - 0}{\sigma_{INB}}$$

2. Estimate the unit normal loss integral (UNLI) for the standardized distance. UNLI (L(D)) is the probability density of D within a standard normal distribution (f(D) minus the product of D and 1 minus the cumulative density of D within a standard normal distribution (F(D)).

$$L(D) = f(D) - D(1 - F(D))$$

The probability density of a standard normal distribution is defined. However, the cumulative density of a standard normal distribution is not solvable in closed form.

---

[30] I am grateful to Jeremy Oakley from the University of Sheffield for statistical advice relating to this section particularly with relation to notation.

Estimates of the cumulative density can be obtained based on quadrature from either statistical tables or using standard statistical or spreadsheet software.

$$f(D) = \frac{e^{-D^2/2}}{\sqrt{2\pi}}$$

$$F(D) = \int_{-\varpi}^{D} \frac{e^{-D^2/2}}{\sqrt{2\pi}} dx$$

3. EVPI can now be estimated as the product of the standard deviation of INB and UNLI for the standardized distance.

$$EVPI = \sigma_{INB} * L(D)$$

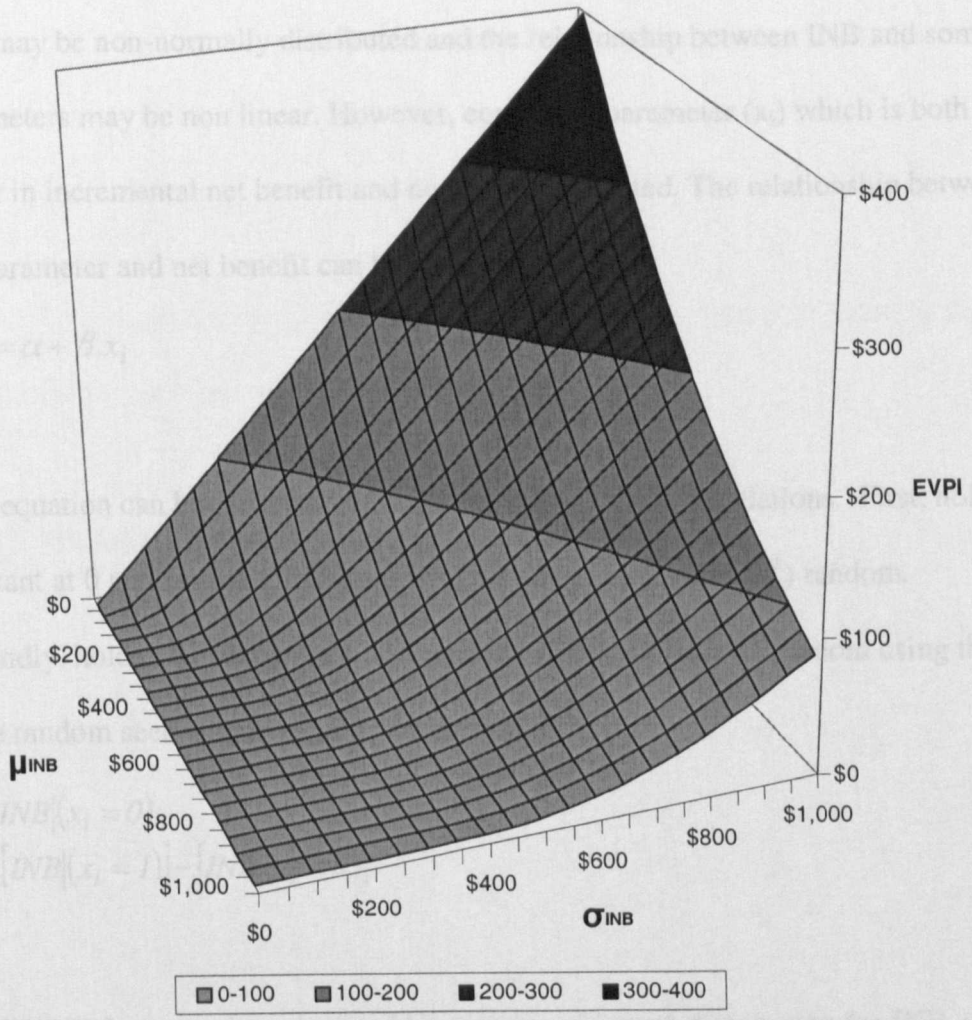
Consider a scenario where INB is normally distributed with an expected value of $1000 with a standard deviation of $1000. EVPI can be calculated as follows.

$$D = \frac{1000 - 0}{1000} = 1$$
$$L(1) = 0.083$$
$$EVPI = \sigma_{INB} * L(D) = \$83$$

Varying both the expected value and the standard deviation of INB confirms two of the fundamental principles relating to EVPI (Chilcott et al. 2003a). First, EVPI will be greater the greater the uncertainty around the mean. Secondly EVPI will be greater the lower the expected value of INB (Figure 4.10).

A recent review of the use of modeling in research prioritization, found no examples of where the EVPPI has been estimated using the unit normal loss integral method (Chilcott et al. 2003a). Recent work in estimating EVPPIs has ignored this method, primarily due to the focus on situations where the distribution of INB is non-normal.

Figure 4.10: Impact of Expected Value and Standard Deviation of INB on Global EVPI



However, in situations when the global EVPI can not be estimated through L(D), it may still be an appropriate method for calculating EVPPIs for parameter(s).

INB may be non-normally distributed and the relationship between INB and some parameters may be non linear. However, consider a parameter (x_i) which is both linear in incremental net benefit and normally distributed. The relationship between the parameter and net benefit can be expressed as

$$INB = \alpha + \beta . x_i$$

This equation can be estimated through two Monte Carlo simulations. First, hold (x_i) constant at 0 and estimate $INB|x_i=0$ with all other parameters (X^c) random. Secondly, hold (x_i) constant at 1 and estimate $INB|x_i=1$ with X^c random using the same random seed as previously. By definition:

$$\alpha = INB|(x_i = 0)$$
$$\beta = [INB|(x_i = 1)] - [INB|(x_i = 0)]$$

The EVPPI for x_i can be calculated by deriving a normal distribution for INB, which represents the uncertainty around INB propagated by x_i given the uncertainty around other parameters.

$$\mu_{INB_{x_i}} = \alpha + \beta * \mu_{x_i}$$
$$\sigma_{INB_{x_i}} = \beta * \sigma_{x_i}$$
$$D_{x_i} = \frac{\mu_{INB_{x_i}}}{\sigma_{INB_{x_i}}}$$
$$EVPPI_{x_i} = \sigma_{INB_{x_i}} * L(D_{x_i})$$

From the above, one can also derive EVPPIs for a sub-set of parameters which have the same desired properties. Consider the situation where X_i represents a sub-set of X of size j in which all parameters are normally distributed and have a linear relationship with INB. The relationship between the parameters and INB can be expressed as:

$$INB_{X_i} = \alpha + \sum_{i=1..j} \beta_i * x_i$$

and can be estimated through j+1 Monte Carlo simulations using the same random seed and alternate fixed values of X_i.

The EVPPI for X_i can now be derived as follows:

$$\mu_{INB_{X_i}} = \alpha + \sum_{i=1...j} \beta_i * \mu_{x_i}$$

$$\sigma_{INB_{X_i}} = \sqrt{\sum_{i=1...j} (\beta * \sigma_{x_i})^2}$$

$$D_{X_i} = \frac{\mu_{INB_{X_i}}}{\sigma_{INB_{X_i}}}$$

$$EVPI_{X_i} = \sigma_{INB_{X_i}} * L(D_{X_i})$$

4.5.3.2 Single MCS Method

When incremental net benefit is not normally distributed, global EVPI can be estimated through the conduct of a single MCS (Felli and Hazen 1998, 1999). This requires the following steps.

1. Conduct a MCS by sampling from the probability density functions for all parameters.

2. Calculate the mean net benefits for each treatment option and identify the optimal option as that with the maximum net benefits.
3. For each replication within the MCS calculate the difference between the net benefits of the optimal treatment and the maximum net benefits across all treatments.
4. Global EVPI is the expected value from step 3.

Felli and Hazen have shown that, if INB is multi-linear in X^c, EVPPI for X_i can be estimated by repeating the steps for estimating Global EVPI with X^c fixed at their expected value[31].

$$EVPPI_{x_i} = E_{X_i}\left[max_t \left\{ NB_t \mid X^c = E\left(X^c\right) \right\} \right] - NB*$$

1. Conduct a MCS by sampling from the probability density functions of the parameters of interest (X_i) with all other parameters fixed at their expected value ($X^c = E(X^c)$).
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.

This method can also act as a proxy for EVPPI if INB is not multi-linear in X^c.

[31] INB is multi-linear in X^c if all parameters within X^c have a linear relationship with INB.

4.5.3.3 Two Stage MCS Method

Brennan, Chilcott and colleagues amongst others have suggested a method of calculating EVPPIs which involve solving both the inner and outer integration through a two stage Monte Carlo simulation (Brennan et al. 2002a, 2002b, Chilcott et al. 2003a). This is conducted as follows:

1. Single values are randomly selected from the probability density functions of the parameters of interest (X_i).
2. The parameters of interest are fixed at the values selected in step 1, and the NB for all treatment options is estimated by conducting MCS by sampling from the probability density functions of all other parameters (X^c).
3. For each simulation conducted in step 2, the net benefit of the optimum therapy from the base analysis is subtracted from the maximum net benefit over all therapeutic options.
4. Steps 1-3 are repeated numerous times with different sets of values for the parameters of interest.
5. EVPPI is then the expectation of values obtained from repeating step 3.

4.5.3.4 Quadrature Method

A second method rooted in the mathematical definition of EVPPI, has been suggested which requires fewer repeat simulations than the two stage MCS method (Coyle et al. 2003b, Coyle et al. 2001b). Instead of a two stage Monte Carlo simulation, the outer integration across the probability density functions of X_i can be achieved through numeric quadrature.

Estimating the EVPPI of x_i would require the following approach:

1. A set of values is determined for the parameter of interest. The values should be equally spaced across the individual's parameters probability function with a high degree of coverage.
2. For each value of the parameter chosen in step 1, the NB for all treatment options is estimated by conducting MCS by sampling from the probability density functions of all other parameters (X^c).
3. For each simulation conducted in step 2, the net benefit of the optimum therapy from the base analysis is subtracted from the maximum net benefit over all therapeutic options.
4. Each estimate from step 3 is weighted by the probability density for the specific value of the parameter.
5. EVPPI is then estimated by integrating across the probability density function using Simpson's rule[32].

Note, that in step 1, the greater the number of values chosen and the higher the degree of coverage the more precise the estimate of EVPPI. In the following sections 101 different values of each parameter are used and values cover at least 99.99% of the probability density function.

The same approach can be used to estimate the EVPPI of X_i of size j.

---

[32] Simpson's rule is a method of numerical quadrature: i.e. it allows estimation of the area under a curve of a specified formula. Simpson's rule requires specifying an upper and lower value for x (x_u and x_l). The interval between x_u and x_l is divided into n smaller intervals of equal length (h): $h = (x_u - x_l)/n$. Under the formula for Simpson's rule, the area under the curve (S) can be defined as:

$$S = \frac{h}{3}.[f(x_l) + 4f(x_{l+h}) + 2f(x_{l+2h}) + 4f(x_{l+3h}) ..... + 2f(x_{l-2h}) + 4f(x_{l+h}) + f(x_u)]$$

1. A set of values is determined for **all** parameters of interest X_i. The values should again be equally spaced across the parameters probability density function with a high degree of coverage.
2. For each possible set of values for the parameters chosen in step 1, the NB for all treatment options is estimated by conducting MCS by sampling from the probability density functions of all other parameters (X^c).
3. For each simulation conducted in step 2, the net benefit of the optimum therapy from the base analysis is subtracted from the maximum net benefit over all therapeutic options.
4. Each estimate from step 3 is weighted by the product of the relevant density of the probability function of one of the parameters of interest.
5. The values obtained from step 4 are integrated across the parameter's probability density function using Simpson's rule.
6. Steps 4 and 5 are repeated for each variable until a final value is obtained.

4.5.3.5 Difference Method

In an evaluation of treatments for Alzheimer disease, Claxton and colleagues adopted an alternative formulation whereby the EVPPI for X_i can be estimated by the difference between global EVPI given uncertainty in all parameters and global EVPI when X_i is fixed (Claxton et al. 2001). Similar methods have been used in an evaluation of management strategies for urinary tract infection (Fenwick et al. 2000). The method suggested by Claxton and others defines EVPPI as follows:

$$EVPPI_{X_i} = EVPI - EVPI|[X_i = E(X_i)]$$

$$EVPPI|[X_i = E(X_i)] = E_{X|X_i}\left[max_t \{NB(t,X)|X_i = E(X_i)\}\right] + max_t\left[E_{X|X_i}\{NB(t,X)|X_i = E(X_i)\}\right]$$

Thus the approach involves the following.

1. Estimate Global EVPI as described in Section 4.5.2.
2. Estimate the NB for all treatment options by conducting MCS by keeping the parameters of interest (X_i) fixed at their expected values and by sampling from the probability density functions of all other parameters (X^c).
3. Calculate the expected value over all replications of the net benefit of the optimum therapy subtracted from the maximum net benefit over all therapeutic options.
4. EVPPI is the difference in the values obtained in step 1 and step 3.

4.5.4 Analysis

Analysis focuses on estimating the EVPPI for each parameter within the decision analysis. The difference, quadrature and two stage MCS methods are used to estimate EVPPI for all parameters within the model. For cost and utility parameters, which are linear in INB and are assumed normally distributed, EVPPI was also estimated by the UNLI method. As INB is linear in all parameters except PPROGRESS, the single MCS method was only an appropriate method for PPROGRESS[33].

[33] Variable definitions are provided in Table 4.1.

In addition to single parameters, EVPPI is estimated for two sets of parameters; utilities (UMILD and USEVERE) and probabilities (PPROGRESS and PIMPROVE).

The number of MCS conducted will affect the accuracy of the predicted EVPPI due to the associated Monte Carlo error. For the base analysis all Monte Carlo simulations involved 5 000 replications. To assess the accuracy of each method with respect to Monte Carlo error, analysis was repeated for a sub set of parameters (utility and transition probabilities) using an extreme number of replications (5 million).

4.5.5 Results

Table 4.3 compares the estimates of EVPPI for each parameter and set of parameters based on the alternative methods[34]. The difference method gave substantially different values from the other methods and can be dismissed as a true measure of EVPPI. The other four methods gave broadly similar values for EVPPI for most parameters. PPROGRESS had the highest EVPPI, followed by utility parameters and PIMPROVE. Cost parameters had little information value.

The values obtained from the UNLI and quadrature methods were more similar than from the two stage MCS method, suggesting that the quadrature method may be a more efficient general method of estimating EVPPI.

[34] Note, that in the conduct of MCS, two specific issues have to be addressed with respect to the uncertainty around individual input parameters; the degree of uncertainty as characterized by the assumed variance and the shape of the uncertainty as characterized by the form of the probability density function. Appendix B addresses the impact on the case study of the latter choice on estimates of INB, global EVPI, EVPPI and the shape of the CEACC. Analysis found that INB and EVPI did not vary substantially by the shape of the probability density function but EVPPI did.

Table 4.3: Estimates of EVPPI Based on Alternative Formulations

	Method of Estimation				
	Single MCS	Difference	Quadrature	Two Stage MCS	UNLI
Single parameters					
PIMPROVE	N/A	28.80	0.69	0.48	N/A
PPROGRESS	5.91	39.88	6.48	6.49	N/A
UMILD	N/A	44.07	2.64	2.49	2.68
USEVERE	N/A	43.56	2.65	2.28	2.68
CCONSM	N/A	3.54	< 0.001	0	< 0.001
CHOSPM	N/A	3.09	< 0.001	0	< 0.001
CADDM	N/A	0.46	< 0.001	0	< 0.001
CCONSS	N/A	33.70	0.23	0.31	0.23
CHOSPS	N/A	23.03	0.04	0.01	0.04
CADDS	N/A	0.22	< 0.001	0	< 0.001
Sets of parameters					
Probabilities	20.21	65.06	24.55	20.63	N/A
Utilities	N/A	77.54	25.36	25.31	25.47

Note: Variable definitions are provided in Table 4.1

Table 4.4 compares the estimates of EVPPI from the four methods based on 5000 and 5 million replications. The values obtained from using 5 million replications differ modestly from analysis based on 5000 replications suggesting that, in this instance, a MCS based on 5000 replications may be sufficient. The results from the UNLI, quadrature and two stage MCS are very similar confirming that each method is estimating the same variable with the difference being error with respect to integral measurement.

4.5.6 Conclusions

The estimation of EVPPI for parameters and sub sets of parameters is an essential component in the analysis to identify the value of obtaining further information given decision making under uncertainty. In addition, EVPPI has been argued to be a theoretically correct measure of the sensitivity of a study's results (Felli and Hazen 1998, 1999). However, to facilitate such usage, EVPPI has to be accurately measured.

In this section, five alternate methods for estimating EVPPI have been identified, described and applied to the case study outlined in Section 3.2. All measures are subject to Monte Carlo error. As the number of replications used to estimate EVPPI increase, appropriate method for the estimation of EVPPI will converge to the same value.

The difference method proposed by Claxton and colleagues is clearly an inappropriate method for estimating EVPPI. The difference method is not rooted in the mathematical definition of EVPPI. However, it had been argued that if the

Table 4.4: Estimates of EVPPI Based on Alternative Formulations and Number of Replications

		Method of Estimation			
		Single MCS	Quadrature	Two Stage MCS	UNLI
PIMPROVE	r = 5000	N/A	0.69	0.48	N/A
	r = 5 million	N/A	0.71	0.56	N/A
PPROGRESS	r = 5000	5.91	6.48	6.49	N/A
	r = 5 million	6.55	6.52	6.55	N/A
UMILD	r = 5000	N/A	2.64	2.49	2.68
	r = 5 million	N/A	2.52	2.52	2.52
USEVERE	r = 5000	N/A	2.65	2.28	2.68
	r = 5 million	N/A	2.52	2.55	2.52

Based on a threshold value for a QALY of $50 000.

Note: Variable definitions are provided in Table 4.1

relationship between a parameter and the outcome of interest is not markedly non linear then the difference method would be a suitable means to estimate EVPPI (Claxton et al 2001). The results from this study dispute this proposition as there is a substantial difference in values obtained from this and the other methods. Thus, this method is demonstrated empirically to be an inappropriate measure of EVPPI.

The single MCS and UNLI methods are computationally efficient methods of estimating EVPPI. However, they are only appropriate for estimating EVPPI in specific limited circumstances relating to the mathematical relationship between input parameters and INB.

The quadrature and two stage MCS methods can be considered general methods for estimating partial EVPPI as they can be applied in all circumstances. The methods are comparable. By increasing the number of MCS used, both methods would return similar values converging to the true value of EVPPI. Both methods are computationally complex. However, based on the methods used in this study the quadrature method may be considered more computationally efficient.

The conclusion reached is that where appropriate EVPPI should be estimated using either the single MCS or UNLI method. When neither of these methods is appropriate, the quadrature method can be used. However, the quadrature method is computationally complex. To reduce the complexity required in estimating EVPPI it may be useful to conduct provisional analysis to identify parameters for which there is likely value in obtaining further information. In the following section, methods for identifying the importance of parameters to the uncertainty around outcomes are

considered as potential means of screening parameters for consideration of a full value of information analysis.

4.6 ALTERNATE METHODS FOR ASSESSING PARAMETER IMPORTANCE

4.6.1 Introduction

In the previous section, calculation of EVPPI was identified as the first stage in estimating the value of information for parameters and sub-sets of parameters. In addition, EVPPI is a theoretically correct measure for determining the importance of parameters in terms of their contribution to the uncertainty around outcomes of interest.

In this section, the concept of importance analysis is introduced and discussed with respect to alternate methods for acting as a screening mechanism[35]. Methods are detailed and applied to the case study described in Section 4.2.

Analysis of the relative importance of input parameters on the reliability of systems is a key aspect of safety analysis (Cheok et al. 1998, Levitin and Lisnianski 1999, Eisenberg and Sagar 2000). Importance analysis involves the use of techniques to determine how different input parameters contribute to the uncertainty over outcomes of interest (Eisenberg and Sagar 2000, Saltelli et al. 2000). Such analyses are common in other disciplines where probabilistic models are used to determine the expected values of outcomes of interest. Techniques used to assess parameter

[35] Much of this section has been published in the Journal of Clinical Epidemiology (Coyle et al. 2003c). I am grateful for comments received from anonymous referees as well as the input of my supervisors.

importance are called importance measures and involve ranking input parameters by their contribution to uncertainty.

Importance analysis differs from standard sensitivity analysis where the aim of the analysis is to address the degree of uncertainty around an outcome measure (Briggs et al. 1994). Instead, the focus of importance analysis is to identify quantitatively those parameters which contribute most to the uncertainty. This can be argued to be more pertinent to decision makers in that it addresses decisions that have to be taken; i.e. given the evidence available what is the most appropriate action and what further information should be collected.

Probabilistic methods such as Monte Carlo simulation techniques have been identified as suitable bases for the conduct of importance analysis, in that the techniques allow for 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 (Magnusson et al. 1996, Saltelli et al. 2000, Helton 1993).

Alternative methods for identification of the importance have been applied to models dealing with uncertainty relating to decisions addressing a wide spectrum of public and private policy issues – e.g. health interventions, nuclear safety, fire safety, radioactive waste (Magnusson et al. 1996, Helton 1993, Eisenberg and Sagar 2000, Hamby 1995, Baker 2002). Few of these measures have been considered in economic analysis. This section reports the results of applying a range of importance measures to the case study detailed in Section 4.2 (Coyle et al. 2003b).

If the focus of analysis is merely to identify those parameters which impact decision uncertainty and not to determine an optimal research strategy then importance measures that are shown to provide a similar ranking of parameters as EVPPI may be useful. However, this is a non optimal approach to analysis as it is contrary to the normative framework adopted. However, given the complexity required in calculating EVPPI, alternate measures of importance may be useful as alternate methods of screening parameters: i.e. identify those parameters contributing most to decision uncertainty for which further analysis of EVPPI is desirable. Techniques must then give similar rankings of parameters and be computationally efficient.

Thus, the objective of this section are as follows:

- To describe alternate measures of parameter importance
- To compare the rankings of parameters based on each technique
- To assess the computational efficiency of each method in terms of the number of MCS required and the need for complex statistical analysis.
- To identify potential measures which may act as a screening method for further VOI analysis

4.6.2 Methods

4.6.2.1 Identification and Classification of Importance Measures

A thorough review of the literature using databases such as Medline and HealthStar; internet free text searching and hand searching of relevant journals and articles relating to health economics, risk assessment, safety analysis and environmental appraisal was conducted to identify potential importance measures.

4.6.2.2 Correlation Coefficients

Use of correlation based measures is one of the most common forms of importance analysis. Many studies have adopted simple correlation coefficients; either Pearson correlation coefficient or Spearman rank correlation (Magnusson et al. 1996, Saltelli et al. 2000, Hertwich et al. 1999, Huijbregts al. 2000).

Spearman rank correlation (R_s) has been shown to be more appropriate than Pearson correlation when non linear relationships between input parameters and outcomes exist which is generally the case with Markov models (Hofer 1999). R_s is simply the Pearson correlation coefficient calculated on rank transformed data (Saltelli et al. 2000, Hamby and Tarantola 1999). For a single input parameter x_i, R_s is defined as:

$$R_s = \frac{\sigma_{r_{x_i}, r_{INB}}}{\sigma_{r\,x_i}\sigma_{r\,INB}}$$

where

$r_{x_i} = x_i$ *rank transformed*

$r_{INB} = INB$ *rank transformed*

R_s can be estimated for all parameters through a single Monte Carlo simulation by the following.

$$\hat{R}_s = 1 - \frac{6D}{R^3 - R}$$

where

$$D = \sum_{n=1}^{R}\left(r_{x_n} - r_{INB_n}\right)^2$$

$n = n^{th}$ *replication*

$R =$ *number of replications*

When there is a high degree of correlation between input parameters it is preferable to use partial measures of correlation. Measures are calculated as above, after eliminating any correlations with other variables (Hofer 1999, Saltelli et al. 2000). In the example used to compare each of the methods, it is assumed that there are no correlations between input parameters so partial measures are not used.

4.6.2.3 Variance Based Measures

Ranking input parameters by their contribution to the variance of the outcome has been a commonly suggested measure of parameter importance (Magnusson 1996, Baker 2002, Cullen 1995, Iman and Helton 1988, Bartell et al. 1986). The contribution to variance (CV) can be estimated through repeated Monte Carlo simulation and is defined as proportion of the variance of the outcome explained by the input parameter (Saltelli et al. 2000).

$$CV(x_i) = \frac{\sigma^2_{INB} - \sigma^2_{INB}|x_i}{\sigma^2_{INB}}$$

This is calculated as follows:

1. The variance of INB is estimated by conducting MCS by sampling from the probability density functions of all parameters (X).
2. A set of values is determined for the parameter of interest. The values should be equally spaced across the individual's parameters probability function with a high degree of coverage.
3. For each value of the parameter chosen in step 1, the variance of INB is estimated by conducting MCS by sampling from the probability density functions of all other parameters (X^c).

4. Each estimate from step 3 is weighted by the probability density for the specific value of the parameter.
5. The variance of INB conditional upon x_i, is then estimated by integrating across the probability density function using Simpson's rule.
6. Contribution to variance is estimated based on the formula.

4.6.2.4 Regression Coefficients

Regression coefficients have been suggested as possible measures of parameter importance (Helton 1993, Cullen 1995, Iman and Helton 1988). Given that the absolute coefficients are a function of the relative magnitude of the parameter, it is necessary first to standardize coefficients. Standardization can be conducted 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 incremental net benefit (Saltelli et al. 2000, Morgan and Henrion 1990).

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

where

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

4.6.2.5 Probability Based Measures

A technique referred to as generalized sensitivity analysis has been used in environmental appraisal (Saltelli et al. 2000, Spear and Hornberger 1980, McKenna and Arnold 1998, James et al. 1996, Choi et al. 1999). The technique compares two

cumulative distribution functions for a single input parameter[36]. Parameter importance is measured by the maximum vertical distance between the two density functions (maximum separation distance (MSD)) defined as:

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

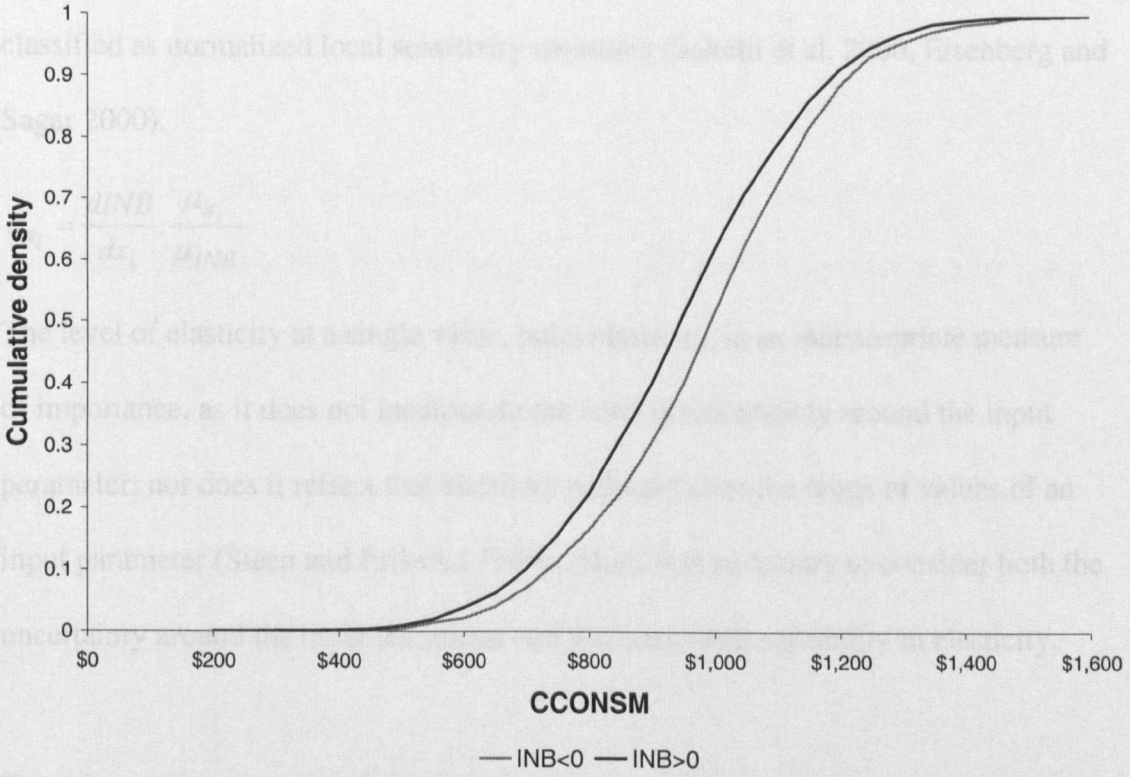
This can be measured from the results of a single MCS sampling from the probability density functions of all parameters (X).

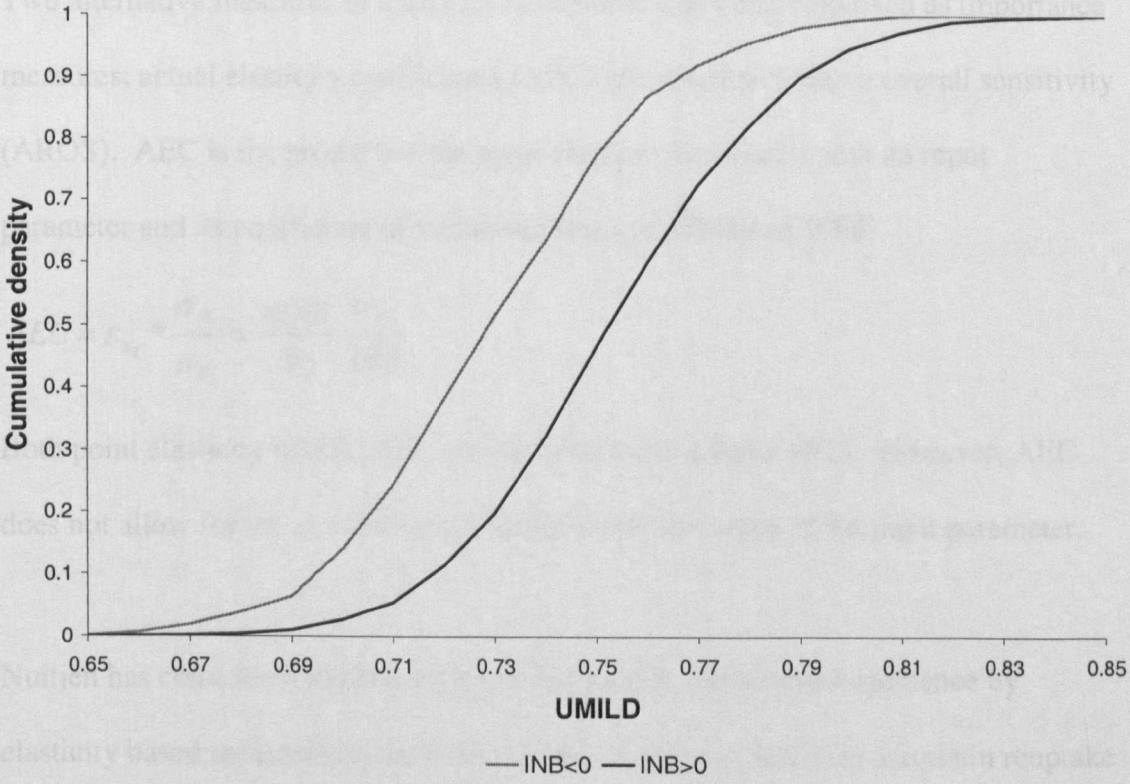
1. Divide the results of the MCS into 2 data sets A and B: where INB >0 and INB<0.
2. For each value (k) within the range of values of x_i, calculate the proportion of A where x_i is less than the value (the cumulative density).
3. Repeat step 2 for B.
4. For each value within the range of values of x_i, calculate the absolute difference between the cumulative densities for A and B.
5. The MSD is the maximum value from step 5.

The difference between the two density functions will be larger for parameters which have a major impact on INB compared to those with minimal impact. Figure 4.11 shows the cumulative density functions dichotomized by INB for UMILD and CCONSM.

[36] Birnbaum importance and Fussell Vesely importance are measures commonly used in environmental appraisals and have been recognized as an appropriate measure of parameter importance for the assessment of safety of nuclear reactors (Birnbaum 1969, Fussell 1975, Cheok et al. 1998). Both techniques are concerned with identifying those components of a system whose failure are most important with respect to system failure. However, both measures require that both input parameters and outcomes are binary. Generalized sensitivity analysis is equivalent to these techniques in the analysis of continuous data.

Figure 4.11: Cumulative Density Functions of UMILD and CCONSM for INB <0 and >0





4.6.2.6 Elasticity Based Measures

Elasticity (ε) is a measure of the change in the value of the outcome to a change in the value of an input parameter (Lipsey et al. 1999). Such measures have also been classified as normalized local sensitivity measures (Saltelli et al. 2000, Eisenberg and Sagar 2000).

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

The level of elasticity at a single value, point elasticity, is an inappropriate measure of importance, as it does not incorporate the level of uncertainty around the input parameter; nor does it reflect that elasticity will vary over the range of values of an input parameter (Steen and Erikstad 1996). Thus, it is necessary to consider both the uncertainty around the input parameter and the associated variability in elasticity.

Two alternative measures of elasticity have previously been employed as importance measures; actual elasticity coefficients (AEC) and absolute relative overall sensitivity (AROS). AEC is the product of the point elasticity associated with an input parameter and its coefficient of variation (Steen and Erikstad 1996).

$$AEC = \varepsilon_{x_i} * \frac{\sigma_{x_i}}{\mu_{x_i}} = \frac{dINB}{dx_i} \cdot \frac{\sigma_{x_i}}{INB}$$

Both point elasticity (ε) and AEC can be calculated without MCS. However, AEC does not allow for the variability in elasticity over the range of the input parameter.

Nuitjen has considered various methods of assessing parameter importance by elasticity based measures within an economic analysis of selective serotonin reuptake inhibitors (SSRI) use (Nuitjen 1999, Nuitjen and Hardens 1997). First, Nuitjen and Hardens (1997) considered two concepts point and range sensitivity. Point

sensitivity is simply the elasticity at the base value for each input parameter. Range sensitivity equals the ratio of the absolute difference in outcome based on the minimum and maximum values of an input parameter and the expected value of the outcome.

Nuitjen (1999) further refined range and point sensitivity by incorporating probability distributions in the calculation of AROS. AROS involves estimating the responsiveness of outcome to values for the input parameters by linear regression analysis. The coefficient from the regression analysis is then used to estimate elasticity across the range of the input parameter and an overall elasticity measure is calculated through quadrature. Thus, $dINB/dx_i$ is assumed fixed for all values of x_i, though ε will vary by x_i.

Although already technically complex, AROS requires the assumption that the relationship between input parameters and outcomes are linear. An alternative approach does not require this restrictive assumption (Coyle et al. 2003b). The elasticity coefficient (EC) for a parameter is defined as the expectation of elasticity over the input parameter.

$$EC(x_i) = E_{x_i}(\varepsilon_{x_i})$$

This can be estimated through multiple MCS and numerical quadrature as follows:

1. A set of values is determined for the parameter of interest. The values should be equally spaced across the individual's parameters probability function with a high degree of coverage.

2. For each value of the parameter chosen in step 1, INB is estimated by conducting MCS by sampling from the probability density functions of all other parameters (X^c).
3. From step 2, ε is calculated
4. Each estimate from step 3 is weighted by the probability density for the specific value of the parameter.
5. EC is then estimated by integrating across the probability density function using Simpson's rule.

4.6.2.7 Entropy

Entropy (H) is a measure of the degree of dispersion of values for an outcome measure (Krzykacz-Hausmann 2001, Kapur 1989). 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. For INB this is defined as :

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

As INB is continuous and need not take any identifiable form of probability density function, probability density has to be estimated empirically. For the example used to illustrate these techniques, the entropy of INB was estimated as follows:

1. INB estimates from the base MCS were grouped into INB categories (INB_c) with a width of \$250 (e.g. $0 < INB < 250$, $250 < INB < 500$,.....).
2. The probability that INB ($p(INB_c)$) will fall into each category is estimated based on the proportion of replications in each category from the base MCS.
3. Entropy is the negative of the sum of the product of $p(INB_c)$ and log ($p(INB_c)$) for all categories where $p(INB_c) > 0$.

The mutual information (MI) or relative entropy between two variables is defined as the difference between the entropy of INB as estimated above and the entropy of INB conditional on x_i ($H(INB|x_i)$) (Felli and Hazen 1998).

$$MI_{x_i} = H(INB) - H(INB|x_i)$$

where

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

$H(INB|x_i)$ can be estimated as follows:

1. A set of values is determined for the parameter of interest. The values should be equally spaced across the individual's parameters probability function with a high degree of coverage.
2. For each value of the parameter chosen in step 1, a MCS is conducted sampling from the probability density functions of all other parameters (X^c).
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.

4.6.3 Assessment of Techniques

The following techniques were applied to the case study: rank correlation, standardized regression coefficient, contribution to variance, maximum separation distance, elasticity coefficient and mutual information. In addition, a proxy measure

of EVPPI through use of the single MCS method is considered for all parameters (not just those linear in net benefit) [37].

Techniques were assessed by the relative importance of each input parameter compared to EVPPI (as measured by the quadrature method) and by their computational efficiency. Comparison of the importance of parameters was by both the rankings of input parameters and a calibrated importance score with a scale of 0 to 100. A score of 0 is the minimum possible score for each technique. A score of 100 is given to the maximum absolute score over all parameters for each technique.

Analytical complexity was assessed in terms of the number of Monte Carlo simulations required, the complexity of data manipulation and the requirement for additional statistical software.

For those methods which were shown to give similar rankings as EVPPI and were computationally efficient, further analysis explored the impact of alternate values of λ on parameter importance. Base analysis related to a value of λ of $50 000.

4.6.4 Results

4.6.4.1 Importance of Individual Parameters

Table 4.5 reports the raw scores for each importance measure for each input parameter. Table 4.6 reports the ranking of each input parameter by each measure. Table 4.7 reports the calibrated scores by each importance measure.

---

[37] For all variables except PPROGRESS the single MCS method gave different values than the quadrature method. This confirms that the single MCS method is only a proxy method of estimation in most cases.

Table 4.5: Raw Scores for Each Input Parameter by Importance Measure

Input Parameter	EVPPI [1]	Proxy EVPPI [2]	Rs	SRC	MSD	EC	MI	CV
PIMPROVE	0.69	0.69	0.35	0.37	0.25	5.01	0.02	0.16
PPROGRESS	6.48	5.91	-0.53	-0.53	0.38	7.20	0.05	0.30
UMILD	2.64	3.18	0.37	0.41	0.32	76.92	0.02	0.14
USEVERE	2.65	2.87	-0.39	-0.42	0.30	41.42	0.02	0.17
CCONSM	0.00	0.00	-0.11	-0.11	0.12	0.55	0.00	0.01
CHOSPM	0.00	0.00	-0.15	-0.13	0.11	0.67	0.00	0.02
CADDM	0.00	0.00	-0.04	-0.03	0.07	0.16	0.00	0.00
CCONSS	0.23	0.42	0.33	0.32	0.29	1.79	0.02	0.11
CHOSPS	0.04	0.04	0.29	0.28	0.23	1.49	0.01	0.09
CADDS	0.00	0.00	0.07	0.07	0.05	0.32	0.00	0.01

[1] EVPPI based on quadrature method
[2] EVPPI based on single MCS method

Note: Variable definitions are provided in Table 4.1

Table 4.6: Rank Ordering of the Importance of Parameters

Input Parameter	EVPPI [1]	Proxy EVPPI [2]	Rs	SRC	MSD	EC	MI	CV
PIMPROVE	4	4	4	4	5	4	3	3
PPROGRESS	1	1	1	1	1	3	1	1
UMILD	3	2	3	3	2	1	4	4
USEVERE	2	3	2	2	3	2	2	2
CCONSM	8	8.5	8	8	7	8	8	8
CHOSPM	7	8.5	7	7	8	7	7	7
CADDM	10	8.5	10	10	9	10	10	10
CCONSS	5	5	5	5	4	5	5	5
CHOSPS	6	6	6	6	6	6	6	6
CADDS	9	8.5	9	9	10	9	9	9

[1] EVPPI based on quadrature method
[2] EVPPI based on single MCS method

Note: Variable definitions are provided in Table 4.1

Table 4.7: Calibrated Importance Scores for Each Input Parameter by Method

Input Parameter	EVPPI [1]	Proxy EVPPI [2]	Rs	SRC	MSD	EC	MI	CV
PIMPROVE	10.6	11.7	65.7	69.3	64.8	6.5	47.1	54.3
PPROGRESS	100	100	100	100	100	9.4	100	100
UMILD	40.7	53.8	70.9	76.8	83.6	100	46.5	48.4
USEVERE	40.9	48.6	74.3	79.3	78.8	53.9	50.8	55.5
CCONSM	0.0	0.0	21.7	19.9	31.1	0.7	2.1	4.4
CHOSPM	0.0	0.0	28.0	24.9	28.8	0.9	4.3	7.7
CADDM	0.0	0.0	7.5	5.8	19.2	0.2	0.0	0.4
CCONSS	3.6	7.1	62.4	61.9	76.0	2.3	33.2	37.5
CHOSPS	0.6	0.7	54.8	53.5	60.4	1.9	25.1	29.8
CADDS	0.0	0.0	13.8	12.7	14.2	0.4	0.0	2.1

[1] EVPPI based on quadrature method
[2] EVPPI based on single MCS method

Note: Variable definitions are provided in Table 4.1

EVPPI, proxy EVPPI, rank correlation and standardized regression coefficients all have a similar ranking with PPROGRESS as the most important parameter followed by the two utility parameters followed by PIMPROVE and then all cost parameters. Maximum separation distances gave a similar ranking for the first three parameters but with different rankings for all other parameters. The other three methods gave different rankings for the first four parameters.

As well as divergences in ranking there are significant divergence in terms of the calibrated importance scores between each method. Rank correlation, standardized regression coefficients and maximum separation distances gave similar calibrated scores. However, none of these methods gave similar calibrated scores to EVPPI. EVPPI and proxy EVPPI gave similar calibrated scores.

4.6.4.2 Complexity

Table 4.8 details the relative complexity of each of the techniques considered. Complexity of EVPPI depends on the appropriate method for each particular input parameter. For this analysis, estimates of EVPPI using the quadrature method were used. Under this method calculation of EVPPI is complex and computationally efficient methods for screening parameters may be desirable.

Of the other methods, rank correlation is less complex requiring only simple analysis of data from the initial simulation which could be conducted easily within an Excel spreadsheet. Maximum separation distances and standardized regression coefficients can also be calculated within Excel though this requires complex data manipulation. Both these techniques are simple if data is transferred to a more advanced statistical

Table 4.8: Complexity of Importance Measures

	Number of Replications Required	Complexity of data manipulation	Additional statistical software*	Replications required for case study
Importance Measures				
Rank correlation	r	Simple	No	5 000
Standardized regression coefficient	r	Simple	Preferable	5 000
Maximum separation distance	r	Simple	Preferable	5 000
Elasticity coefficient	r*m*k	Complex	No	5 050 000
Partial contribution to variance	r*m*k +1	Complex	No	5 050 001
Expected reduction in entropy	r*m*k +1	Complex	No	5 050 001
EVPPI				
UNLI	2*r*k	Complex	No	100 000
Single MCS	r*k	Simple	No	50 000
Difference method	2*r*k	Simple	No	100 000
Two stage MCS	r*r*k	Simple	No	250 000 000
Quadrature method	r*m*k	Complex	No	5 050 000

k = uncertain parameters (for case study k =10)
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)

*Additional to Excel or other spreadsheet based software

package such as SPSS. Other techniques require repeat Monte Carlo simulations. Contribution to variance, elasticity and entropy require calculating expected values over each input parameter. This involves a significant number of repeat Monte Carlo simulations using a range of fixed values of each input parameter. These techniques also require an extensive degree of data manipulation after the conduct of the simulations.

4.6.4.3 Impact of λ

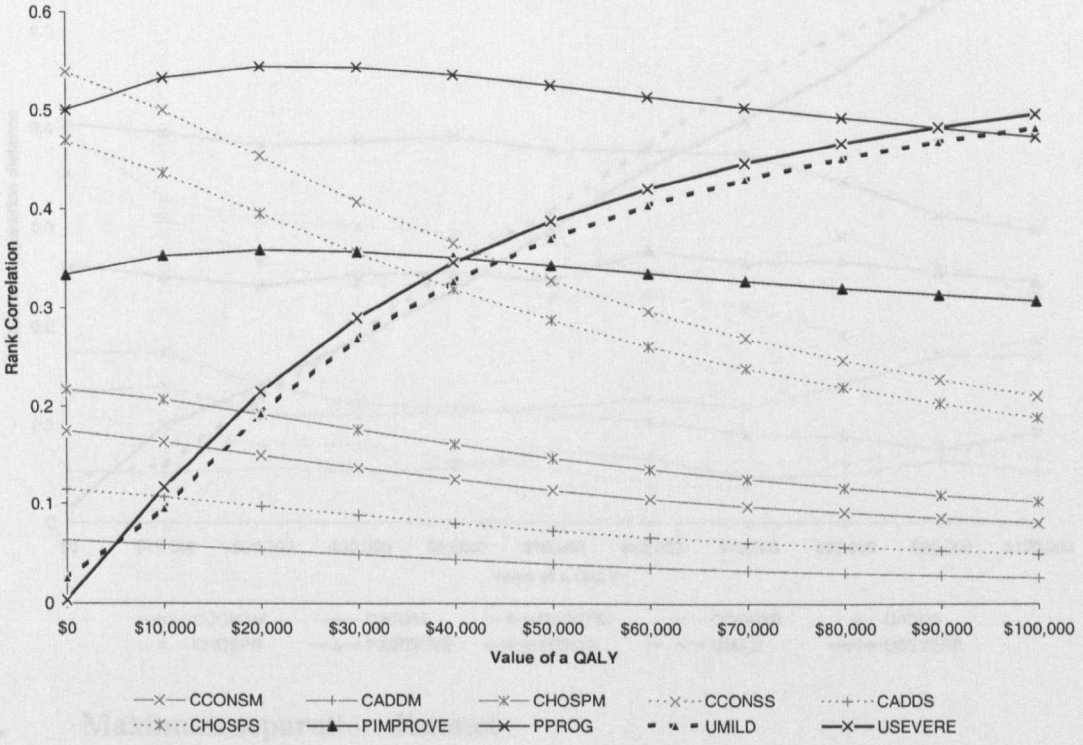
Given the relative values from the importance measures and their relative complexity, elasticity coefficients, partial contribution to variance and entropy are unlikely to be efficient methods for screening variables for estimation of EVPPI.

Figure 4.12 demonstrates the variability in both rankings and values for the following measures over a range of threshold values: EVPPI, proxy EVPPI, rank correlation, standardized regression coefficients and maximum separation distances.

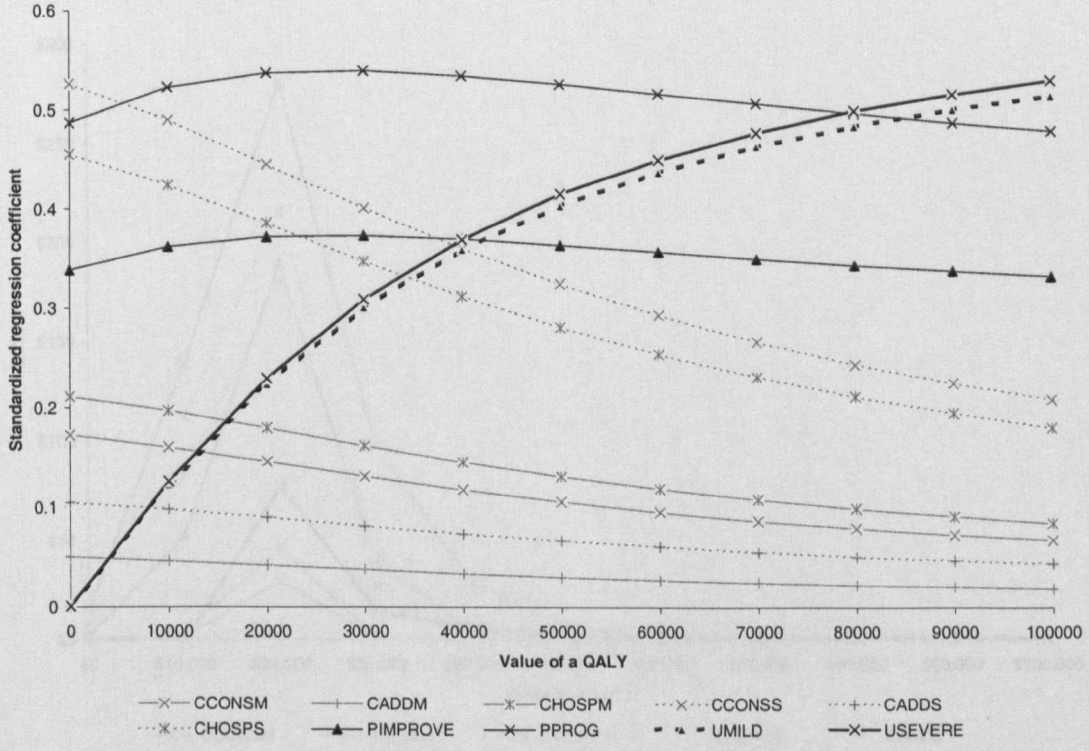
As expected, for all input parameters the raw values for EVPPI peak at the value of the ICER. The ranking of parameters by EVPPI varies by λ with cost parameters being relatively more important the lower the value of λ. Similar findings occur with respect to analysis using the single MCS method for estimating a proxy EVPPI.

For all other methods, the ranking of each parameter varies by λ similarly to EVPPI. However, none of the methods demonstrate a similar peak in relative importance around the ICER.

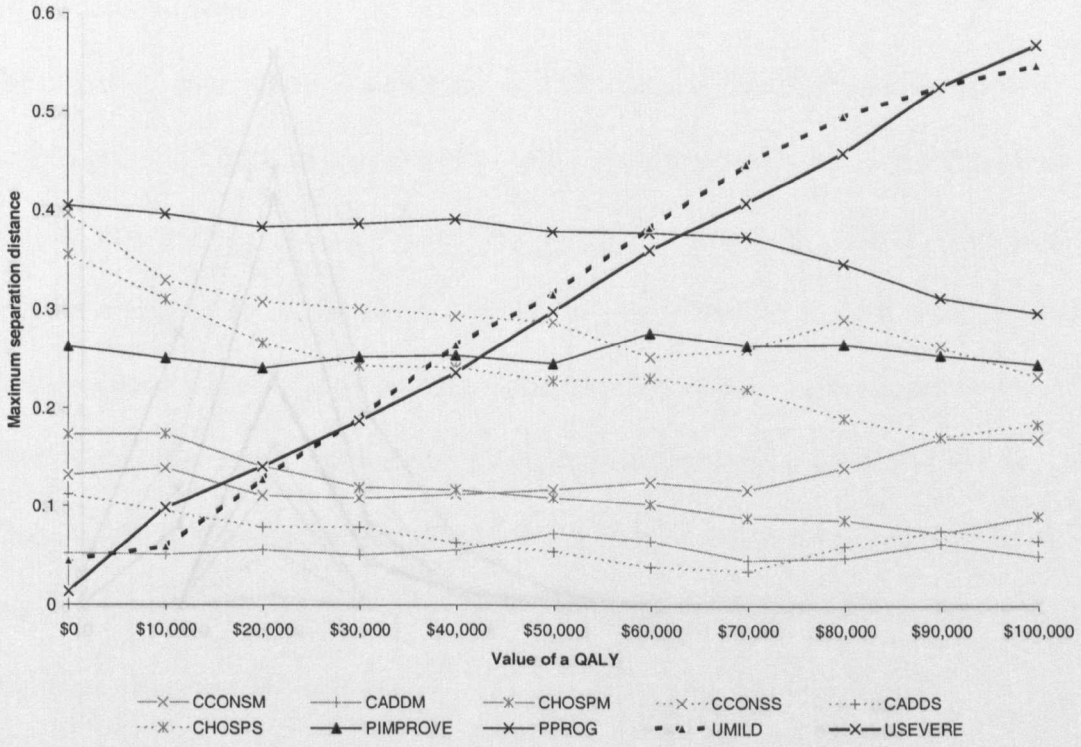
Figure 4.12: Responsiveness of Importance Measures to Changes in λ



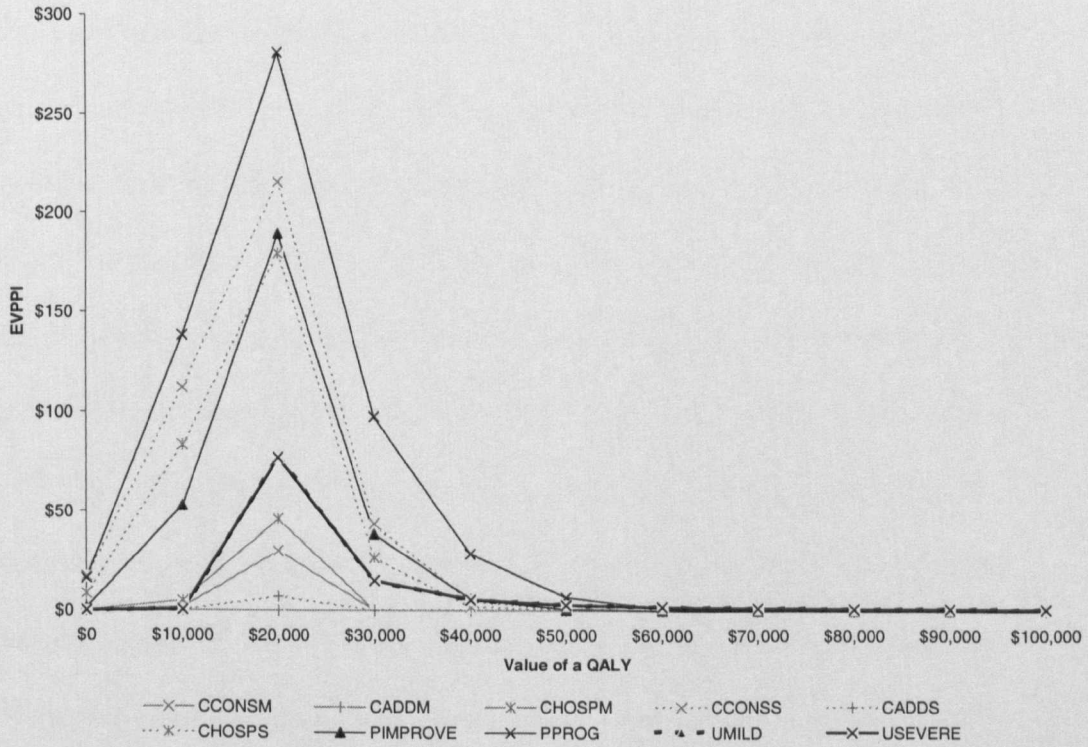
a. Rank Correlation



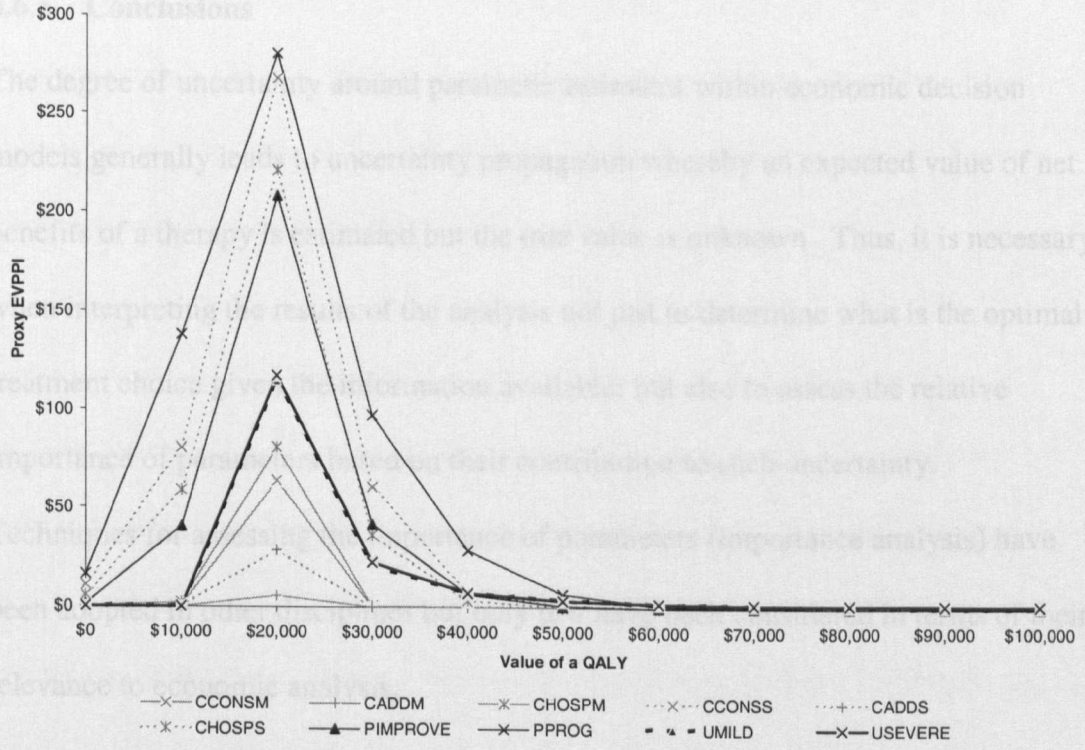
b. Standardized regression coefficient



c. Maximum separation distance



d. EVPPI



e. Proxy EVPPI

4.6.5 Conclusions

The degree of uncertainty around parameter estimates within economic decision models generally leads to uncertainty propagation whereby an expected value of net benefits of a therapy is estimated but the true value is unknown. Thus, it is necessary when interpreting the results of the analysis not just to determine what is the optimal treatment choice given the information available; but also to assess the relative importance of parameters based on their contribution to such uncertainty. Techniques for assessing the importance of parameters (importance analysis) have been adopted in other disciplines but only few have been considered in terms of their relevance to economic analysis.

The appropriateness of each measure of importance needs to be assessed in terms of how results can be interpreted. Such appropriateness must be assessed in the context of the problem facing the decision maker. If it is accepted that the uncertainty that we are interested in is the probability of making an incorrect decision and the consequences of such a decision, then measures of importance that focus on the dispersion of the outcome measure may be inappropriate. 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. Similarly, for most of the techniques, there is difficulty in interpreting what a specific score means in terms of the need for obtaining further information. Thus, the importance scores and rankings do not necessarily help in determining which further information would be the most efficient use of scarce resources. EVPPI is the only technique which directly considers the value from further information.

EVPPI requires intricate analysis which may be computationally inefficient for complex models. Hence, methods of assessing parameter importance which are more efficient may be useful as a screening mechanism for identifying parameters for which computation of EVPPI is warranted.

In the second part of this chapter, several different techniques for importance analysis were identified and applied to an economic model to determine their suitability as a screening mechanism. The measures differ in terms of how they measure the uncertainty in the outcome of interest. Correlation coefficients, regression coefficients, measures of variance and entropy focus on the degree that the input variable is responsible for the dispersion in the outcome variable. Elasticity looks at the responsiveness of outcome to changes in the value of an input parameter. Maximum separation distances focus on the uncertainty concerning the optimal decision based on the outcome measure. Thus, differences in rankings should be expected and the measures which were closest in terms of importance scores were generally similar in how they characterized uncertainty.

Analysis confirms that the importance of specific input parameters varies by a decision maker's threshold value of a QALY. It is clear that both the values and ranking obtained from importance will vary by the threshold employed. Thus, the calculation of EVPPI becomes more complex given both the lack of transparency and the lack of consistency in decision making relating to a threshold value. Analysis must be presented over a range of potential thresholds to allow determination of the optimal research design.

Measures differ in terms of their analytical complexity. Some techniques are relatively simple requiring only the analysis of a base Monte Carlo simulation. However, other techniques require multiple repeat simulations. Analytical complexity should not be considered a hindrance to the use of appropriate techniques. However, the need to conduct analyses for a wide range of threshold values suggests that simple methods of assessing importance may be of use in refining the scale of analysis required with respect to determination of EVPPI and subsequently the expected value of sample information (EVSI). That is, simple techniques, which return relative values consistent with EVPPI, may be used to identify a sub-set of parameters for the conduct of a full analysis.

For a technique to be an appropriate method for identifying parameters for which a full value of information analysis would be appropriate, the technique must provide relative rankings consistent with EVPPI and be computationally efficient. The last criteria effectively eliminates elasticity coefficients, partial contribution to variance and entropy from consideration. Maximum separation distances, rank correlation and standardized regression coefficients reported similar parameter rankings to EVPPI. However, for these measures there were substantive differences in the calibrated importance score for each parameter compared to EVPPI. Furthermore, none of the measures showed a similar peak in importance around the ICER. Given that the relative importance of parameter uncertainty should be greater the more uncertain the optimal decision, this is a major concern with regards to the relative value of these methods as screening mechanisms.

A proxy measure of EVPPI based on the single MCS method provided scores similar to EVPPI. This method is computationally efficient and appears to be the most appropriate method for identifying parameters for which formal calculation of EVPPI is required.

In conclusion, EVPPI is the theoretically correct method for determining parameter importance and is a necessary component to the determination of an efficient research design. The other measures identified have neither an equally appropriate theoretical basis nor practical application. 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. The single MCS method for estimating EVPPI is the most appropriate screening mechanism.

4.7 DETERMINING VALUE OF SAMPLE INFORMATION AND OPTIMAL SAMPLE SIZE

4.7.1 Introduction

In previous sections in this chapter the methods for determining EVPI and EVPPI have been discussed. Following from this, it is necessary to illustrate how such concepts can facilitate determination of the optimal sample size for further studies.

The traditional basis for determining sample size for studies comparing two groups is the avoidance of type 1 and type 2 error. Type 1 error is the probability of incorrectly concluding that there is a difference between groups when no difference exists. Type 2 error is the probability of concluding there is no difference between

groups when one does exist. In determining sample size based on avoidance of such errors, it is necessary to specify two parameters: α (the maximum acceptable Type 1 error) and β (the maximum acceptable type 2 error). Sample size is thus a function of the desired level of power (1- β) and significance (1- α).

An optimal approach to determining sample size would be to maximize the return on the investment in further information. Thus, there are several problems with the traditional basis of determining sample size. Traditional sample size calculations ignore the following factors which can be seen as determinants of the return from generating further information:

- The level of available information : the greater the available information the less further information required
- The potential population affected by the intervention: the greater the potential population the more value from further information
- The incremental cost of the intervention: the greater the incremental cost the less value in obtaining further information
- The costs of obtaining sample information: the greater the cost of sampling the less value in obtaining further information
- Value of outcomes: the greater the value placed on the outcomes of treatment the greater the value of sample information
- Lifetime of intervention: the longer the new information will be useful in determining optimal treatment choices, the greater the value of information
- Uptake of intervention: the greater the probability that physicians or patients will choose to adopt the treatments considered, the more value in further information

An alternate level of error is Type 3 error: the probability of adopting one treatment when another treatment should be preferred. EVPI is simply the product of the probability of type 3 error and the consequences of such error. In the following section the link between EVPI, EVPPI and the expected value of sample information (EVSI) is established and methods of determining optimal sample size for further research illustrated by application to the case study summarized in Section 3.2.

4.7.2 Expected Value of Sample Information

4.7.2.1 Definition of EVSI and MVSI

EVPPI is the opportunity cost of making decision based on the current uncertainty concerning a specific parameter or sub set of parameters. Collecting further sample information can reduce such uncertainty and thus reduce the opportunity cost. The reduction in opportunity cost can be referred to as the expected value of the sample information (EVSI). Thus, the EVSI of a sample with size **s** ($EVSI_s$) is defined as:

$$EVSI_s = EVPPI - EVPPI|s$$

The expected value of sample information will increase as the sample size increases and will tend towards EVPPI as the sample size tends towards infinity. The marginal value of sample information (MVSI) can be defined as the increment to EVSI for one additional sample:

$$MVSI = EVSI_{s+1} - EVSI_s$$

MVSI will tend towards 0 as sample size tends towards infinity.

4.7.2.2 Estimation of EVSI

The expected value of sample information can be estimated as follows (Ades et al. 2004):

1. Estimate EVPPI based on current data
2. For a sample size *s* simulate data collection based on current knowledge
3. Update probability distribution(s) for input parameter(s) by combining prior (original) data with simulated sample data
4. Estimate EVPPI|*s* based on updated distribution
5. Repeat steps 1 – 4 a number of times (5000)
6. Estimate expected value of EVPPI|s
7. EVSI is the result of the subtraction of the value from step 1 and the value from step 6
8. Repeat steps 1 – 7 for various *s*.

EVPPI can be estimated using methods outlined in Section 4.5.3. Estimates of EVPPI|s require combining prior data with simulated data to obtain posterior probability density function(s). In certain instances updating of density functions are straight forward as data are conjugate: i.e. the prior data and the data from the research study can be combined[38]. In other instances updating is more complex and requires the use of specific software such as WinBugs. Within the analysis of the case study, simple updating was possible due to the use of conjugate distributions.

[38] Conjugacy requires that a posterior distribution for the parameter can be obtained by both the prior data and the data from the research study using simple mathematical calculations. Methods for updating density functions follow the methods outlined by Ades and colleagues (2004).

4.7.3 Determining Optimal Sample Size

In the previous section, methods of estimating the EVSI per patient were outlined. By weighting EVSI by the discounted incremental number of patients (n) who will receive treatment based on the new decision we can determine population EVSI. The optimal sample size (s*) for the study will be the point where the return on sampling is the greatest. This will be where the difference between population EVSI and the cost of sample information (CS) is greatest. This is an equivalent condition to where the marginal cost of sample information (MC) is equal to MVSI.

$$s_i^* = max_s(EVSI * n - CS) = s|(MVSI * n = MC)$$

4.7.4 Application to Case Study

4.7.4.1 Background

In Section 3.2, the economic analysis of entacapone in the treatment of Parkinson's disease was summarised. In Section 3.4, the methods for estimating EVPPI were applied to the data used in this study. The case study is now used to demonstrate the methods for estimating EVSI .

4.7.4.2 Potential Study Designs

Three potential study designs are considered. First a costing study could be conducted where by Parkinson's patients with mild and severe disease are identified and their resource use is monitored over a 6 month period. A utility elicitation study could be conducted whereby patients with mild or severe Parkinson's disease are recruited and are asked to conduct a utility exercise. Finally, a clinical cohort study can be conducted whereby a proportion of patients with severe Parkinson's disease are given entacapone. These could be followed up at six months to determine the

proportion improving to mild disease. At 12 months, those who had improved would be followed up to see the proportion progressing to severe disease[39].

The cost of each potential study must be considered based on similar studies conducted previously. For the cost study, the cost per patient recruited is assumed to be $100. The same cost is assumed for the utility study. For the more complex clinical cohort study a cost of $300 per patient enrolled is assumed[40].

In addition to the cost of the studies, before estimating the optimal sample size for potential studies, it is necessary to estimate the use of the treatment : i.e. the patient population which is subject to the opportunity cost of making the decision under uncertainty discounted to present values (n)[41]. This is a function of the incidence of the disease in question (n) (i.e. the potential annual population), the percentage of incidence cases which will receive the treatment (u) (the uptake of treatment), the length of time the treatment will be used before being replaced by a newer treatment, the lifespan of the research (ls)[42] and the relevant discount rate (r):

$$pop = \sum_{t=0,..T} \frac{u * n}{(1+r)^{ls}}$$

The annual incidence of severe Parkinson's disease in Canada is approximately 2600 cases. It is assumed that 50% of patients will receive entacapone and that it will be

---

[39] For the cost and utility studies, the UNLI method is used to estimate EVPI. For the clinical cohort study the quadrature method is used.

[40] In reality, the cost of research is more complex in that it will include both fixed and variable costs. In chapter 8, a more complex model for the costs of research are considered.

[41] Note, that in this and later examples the variables used to determine the cost of sample information and the patient population affected are assumed fixed. This is unlikely and these parameters should be considered uncertain and probability density functions could be specified. Thus, the optimal sample size will be the expected value of s* based on a MCS sampling from these distributions. This is considered further in Section 8.4.

[42] Note, that for the case study it is assumed that research results can be obtained instantaneously. However, the time required to conduct research will limit the lifespan. This is considered in chapter 8.

given to patients for a period of five years before being replaced by a new product. Based on an annual discount rate of 5% the discounted incremental number of patients is 5910.

4.7.4.3 Optimal Sample Size

Figure 4.13 demonstrates how EVSI for each of the potential study designs increases as s increases. The optimal sample sizes for the studies are 28 for the costing study, 170 for the utility study and 200 for the cohort study (Figure 4.14).

Sensitivity analysis was conducted to determine the responsiveness in optimal sample size for the utility study to changes in the annual incidence of disease, the uptake of treatment, the life span of treatment and the cost of sampling. The relationship between parameters and the optimal sample size is non linear (Figure 4.15). Results seem particularly sensitive to assumptions concerning the uptake of research information and the annual incidence of patients. Thus, if there is uncertainty around these parameters further MCS is warranted.

4.8 CONCLUSIONS

Most previous recommendations with respect to handling uncertainty within economic analysis have tended to focus on the conduct of sensitivity analysis which focuses on the robustness of a study's results to changes in parameter values. Such an approach ignores the underlying decision uncertainty which a decision maker must contend with.

Figure 4.13: EVSI for Potential Studies by Sample Size

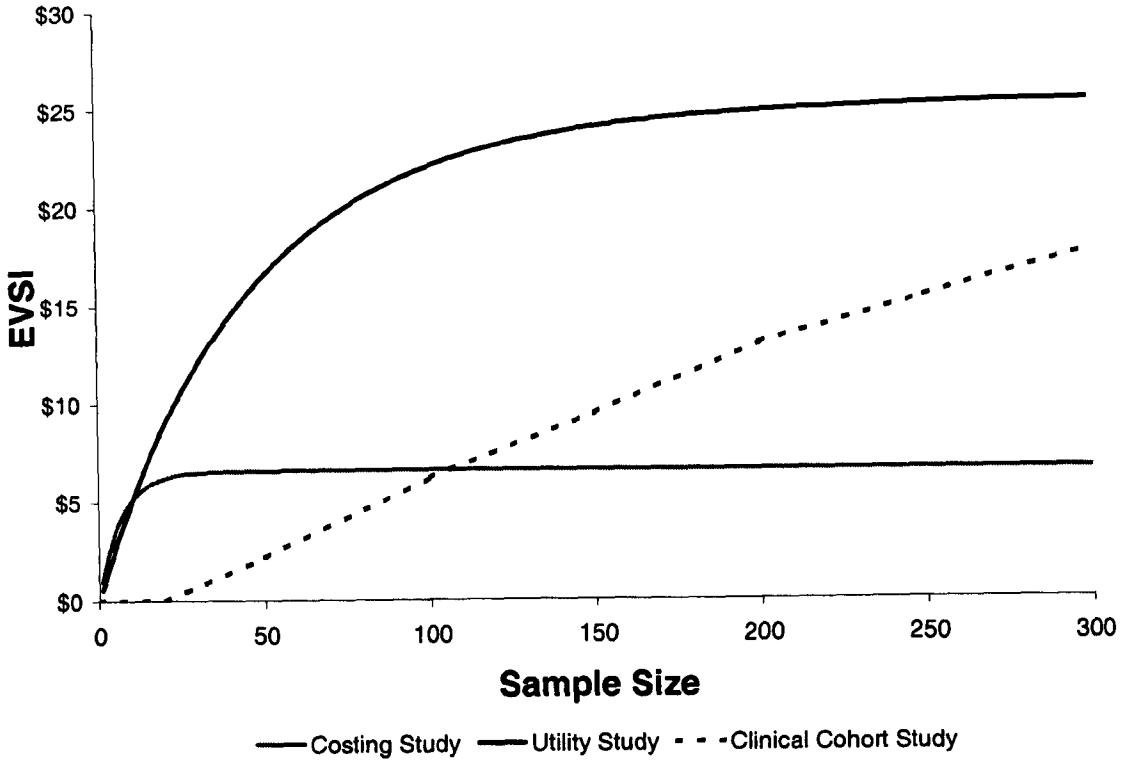
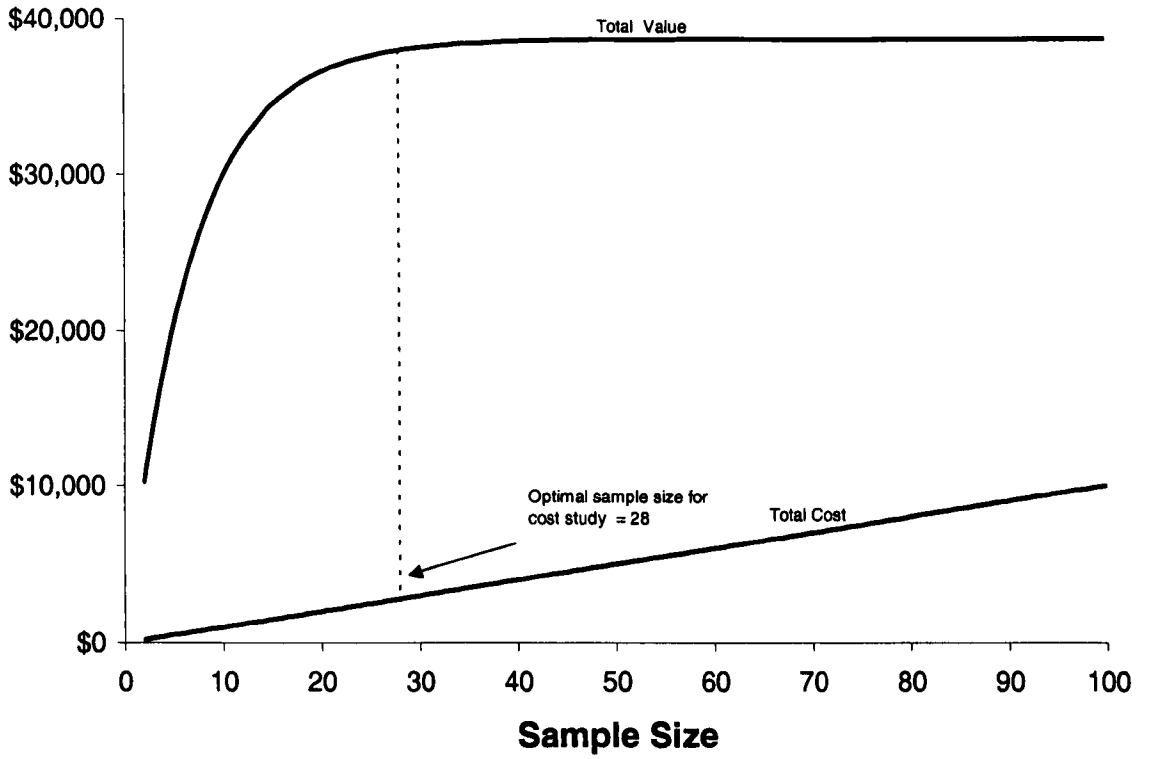
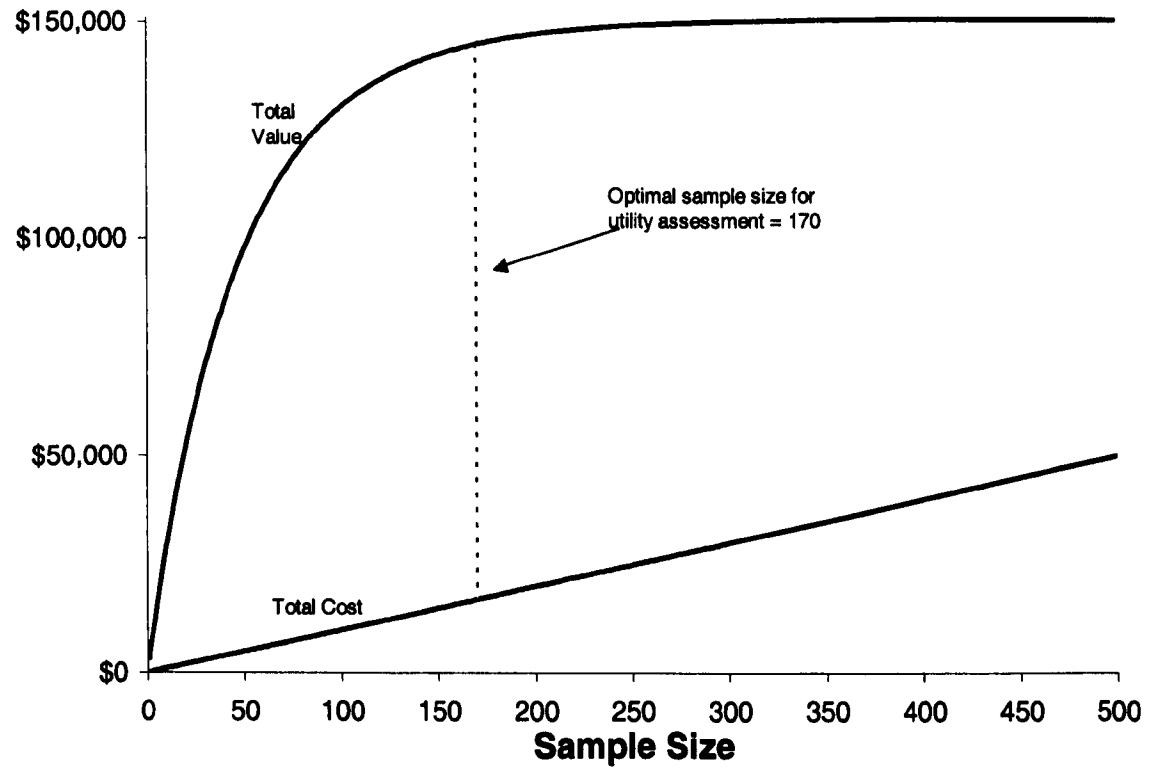


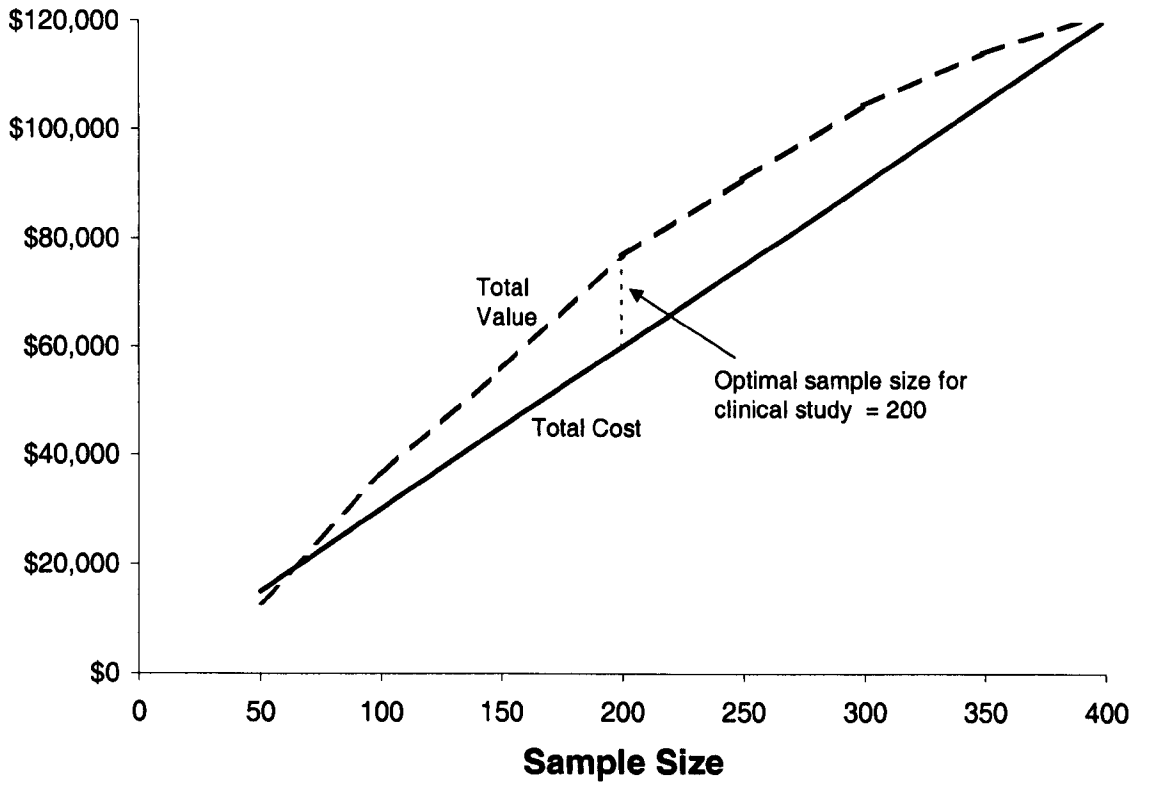
Figure 4.14: Optimal Sample Size for Potential Studies



a. Costing Study

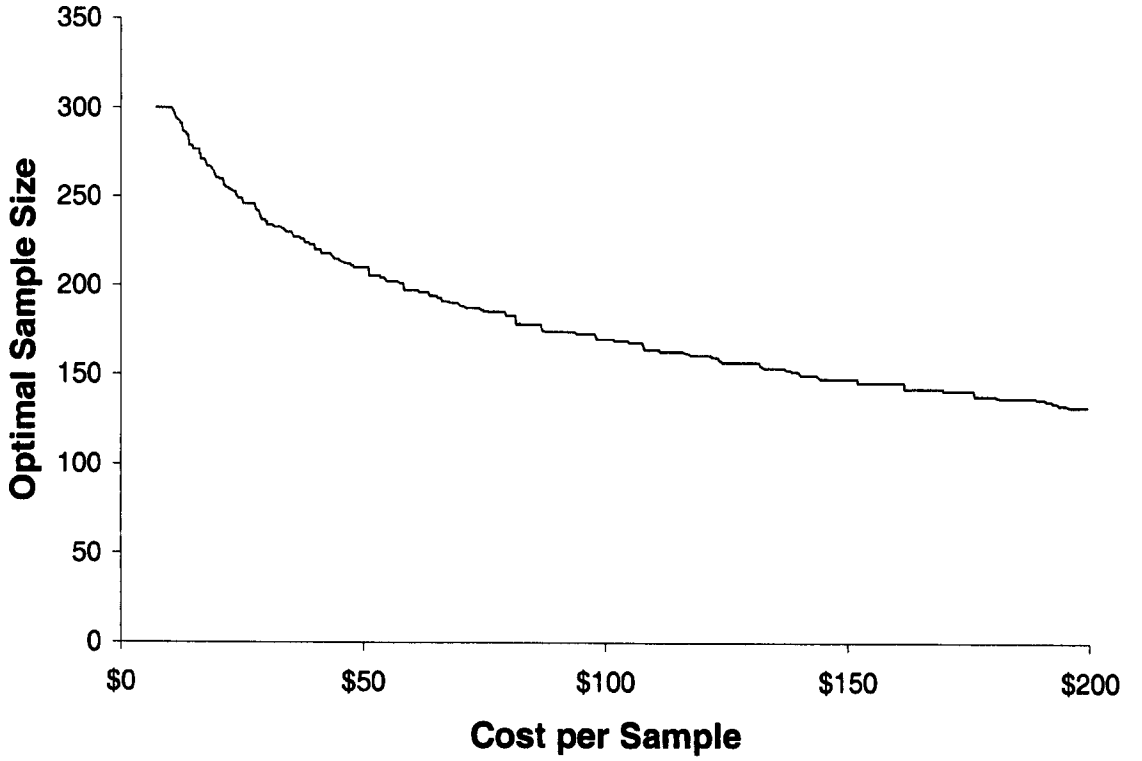


b. Utility Study

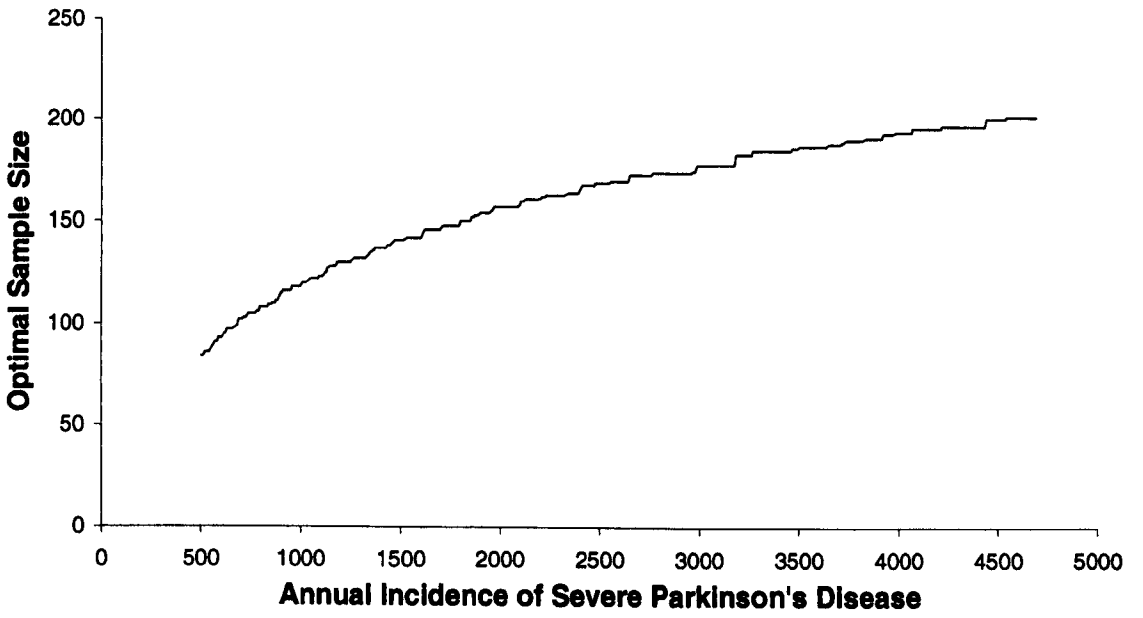


c. Clinical Cohort Study

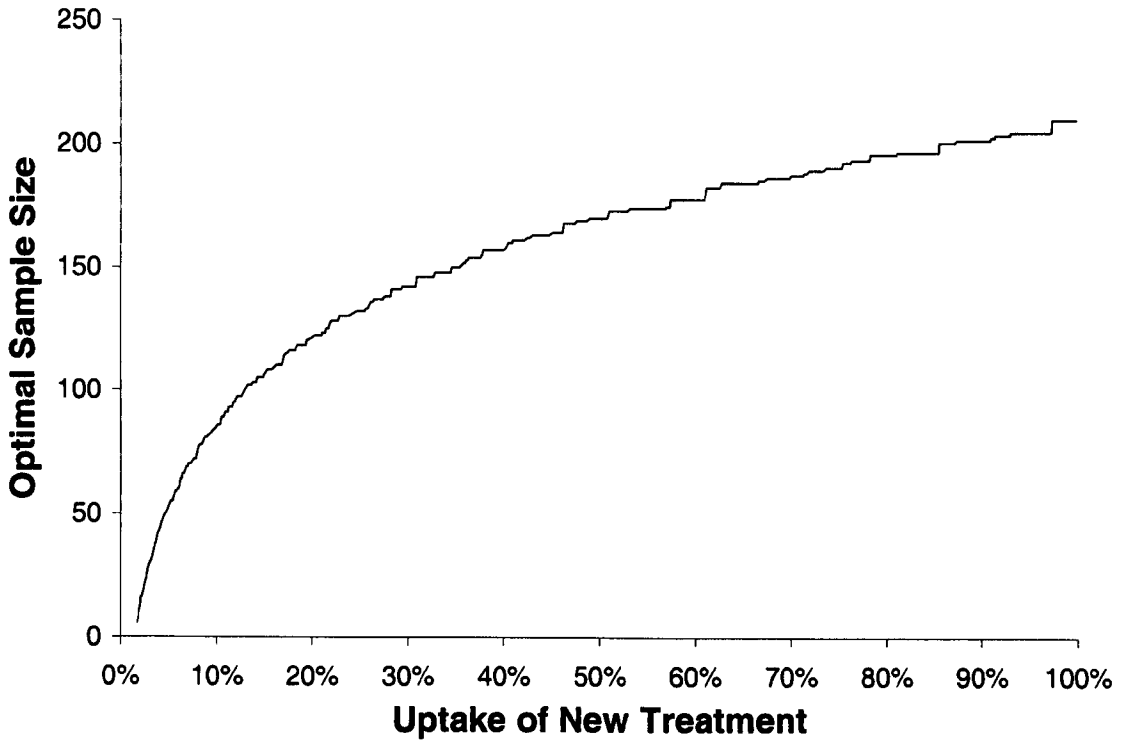
Figure 4.15: Relationship between Parameters and Optimal Sample Size for the Utility Study



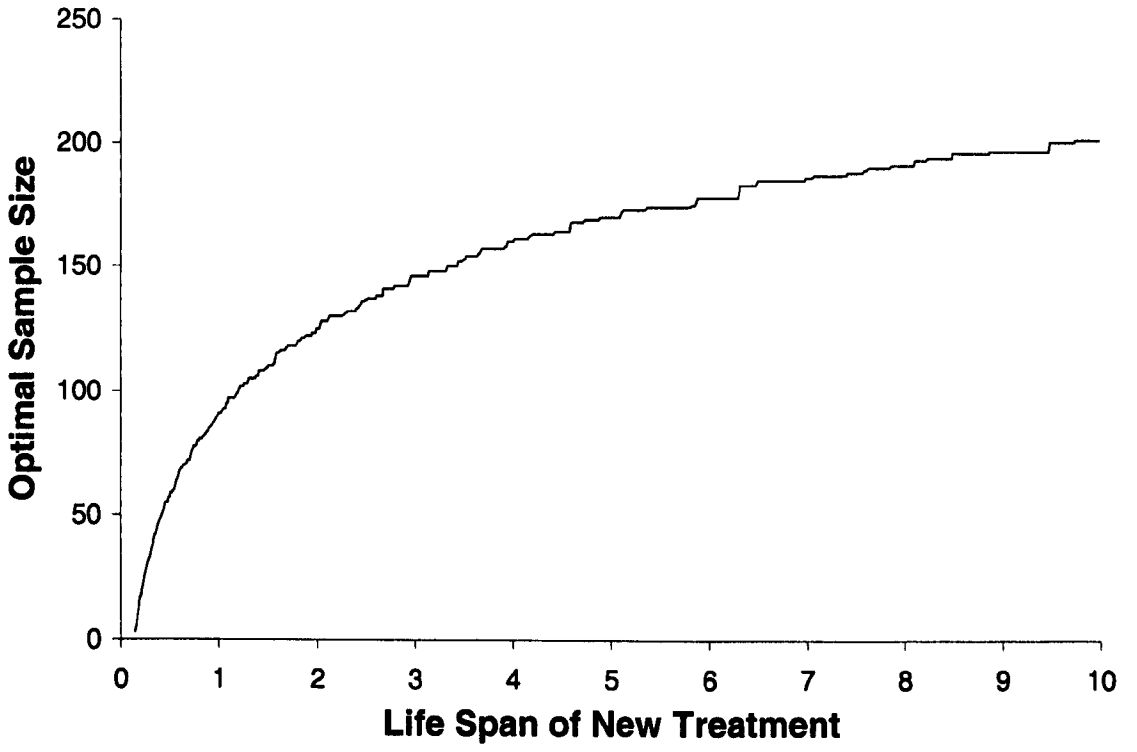
a. Cost per Sample



b. Annual incidence of severe Parkinson's disease



c. Uptake of new treatment



d. Life span of new treatment

In Chapter 2, it was argued that decision makers should solely be concerned with the expected value of outcomes when basing decisions over the optimal treatment choice. Thus, in considering uncertainty it is necessary to obtain precise measures of expected values and the recommended approach is that of probabilistic analysis based on Monte Carlo simulation. This does not however mean that further consideration of the uncertainty over a decision is irrelevant. Rather further focus when handling uncertainty should be on what further information it is justified to collect.

The focus of this chapter has been on conducting a full analysis of uncertainty pertaining to a straightforward economic analysis of treatment for Parkinson's disease. In Section 4.2, the case study was introduced. In Section 4.3, appropriate methods of determining expected values of outcomes related to cost effectiveness were presented. In addition, methods for presenting the level of uncertainty over cost effectiveness are described, applied to the case study and evaluated in terms of their contribution to optimal decision making.

The rest of the chapter relates to determination of an optimal research plan given the uncertainty over the value of therapy within the case study. In Section 4.4, appropriate methods for estimation the value of information were identified. Of the five measures considered, two are appropriate for all input parameters, two are appropriate in specific circumstances and a further measure is wshown to be inappropriate. Given the complexity of certain measures, the focus of Section 4.5 was to identify potential screening measures to simplify the process with respect to identifying an optimal research plan. Although certain measures performed

adequately, the single MCS method for estimating EVPPI was shown to be the most suitable screening measure for identifying parameters with information value. In Section 4.6, an optimal research plan for the case study was identified.

The chapter provides a full description of a normative framework for consideration of uncertainty. First, methods for assessing cost effectiveness given uncertainty are described and applied. In determining the optimal treatment choices for specific patients, uncertainty around input parameters should be characterised by probability density functions, with the expected value of outcomes of interest estimated through Monte Carlo simulation. Secondly, a framework for assessing the value of further information is derived and applied. The framework requires estimation of the expected value of perfect information, the expected value of sample information and the potential cost of future research projects.

The remainder of this thesis is concerned with the application of the framework for handling both variability and uncertainty to a complex economic analysis of treatments for osteoporotic women. Chapters 7 and 8 report analysis where consideration of variability and uncertainty are considered simultaneously to allow identification of both an appropriate limited use strategy for osteoporotic medications as well as an optimal research plan. Before this, it is necessary first to identify appropriate methods for the conduct of economic analysis in osteoporosis (Chapter 5) and secondly to describe in detail the design of the economic model of osteoporosis used in this analysis (Chapter 6).

Chapter 5.

Economic Evaluation for Treatments to Prevent Osteoporotic Fractures

---

5.1 INTRODUCTION

The objective of Chapters 6, 7 and 8 of this thesis is to conduct a full economic analysis relating to alternative treatment options for osteoporosis based on the normative framework outlined in Chapters 3 and 4. The focus of the analysis is to determine which treatments are optimal for which patients (an analysis of variability) and to determine what further research is optimal (an analysis of uncertainty). In Chapter 6, a decision model developed to conduct the analysis is described in detail. In Chapter 7, a stratified cost utility analysis is conducted whereby the optimal treatment choice is identified for different cohorts of patients. Finally, in Chapter 8, a value of information analysis is conducted to ascertain which further research projects are worthwhile.

In this chapter, the focus is on issues relating to the conduct of economic evaluation for treatments to prevent osteoporotic fracture. Section 5.2 contains a brief background to osteoporosis and the available treatment options to prevent fractures. Section 5.3 contains a review of recent guidelines related to the conduct of economic evaluations relating to osteoporosis. Section 5.4 provides a detailed review of existing evaluations of the cost effectiveness of osteoporosis treatments. The review

of existing guidelines provides guidance on how to conduct the required study and, in addition, can be used to determine the quality of existing studies.

5.2 BACKGROUND TO OSTEOPOROSIS

Osteoporosis is a systemic skeletal disease characterized by low bone mass, micro-architectural deterioration of bone tissue, leading to increased bone fragility and a consequent increase in fracture risk (Consensus Development Conference 1993). The most common clinical manifestations of osteoporosis are fractures of the hip, vertebrae or wrist. Approximately 30% of postmenopausal females have osteoporosis according to the World Health Organization (WHO) definition of osteoporosis (Kanis et al. 1994, WHO Study Group 1994). There is evidence of excess mortality associated with a hip fracture which has been estimated to be around 20% (Cooper 1993). However, mortality post fracture is a function of age with evidence of an exponential increase (Papadimitropoulos et al. 1997).

A Canadian study estimated that the health care costs associated with osteoporosis in Canada were $465 million with an additional $563 million for long term care facilities and $279 million for chronic care hospitals (Goeree et al.1996). Thus, therapies which reduce the risk of sequelae associated with osteoporosis may be attractive in that their costs may be partly offset by reducing this burden.

For rheumatologists, osteoporosis is an important disease given that many patients are at increased risk of developing osteoporosis due to the effects of inflammation from their disease, and use of corticosteroids. The focus of treatments for osteoporosis is the prevention of fractures. There are now more therapeutic options

available to treat osteoporosis and many agents are under investigation. The potential costs of drug therapy for the prevention and treatment of osteoporosis represent a considerable burden to those who are paying for them - consumers, third party payers, governments.

The major drug class used in Canada for the prevention of osteoporotic fractures in postmenopausal women are bisphosphonates[43, 44]. Bisphosphonates act by changing the balance between osteoblasts (bone building cells) and osteoclasts (bone-eroding cells). Bisphosphonates slow down the osteoclasts by binding to the surface of bones allowing the osteoblasts to work more effectively. This results in increasing bone mass (density) and thus reduces the risk of fractures (Fleisch 1993).

Three different bisphosphonates are licensed for the treatment of osteoporosis in Canada: etidronate, alendronate and risedronate (Canadian Pharmacists Association 2001). The typical daily dose for alendronate and risedronate are 10 mg and 5 mg respectively. Etidronate is taken cyclically; in that a 400 mg tablet is taken every day for two weeks (14 days) followed by two-and-a-half months (76 days) of calcium supplements (500 mg). The calcium supplements are included in the prescription package. This cycle is repeated four times annually. All drugs should be taken with

---

[43] Focus on postmenopausal women is justified on three counts. The incidence of osteopororis and the risk of fracture for women with osteoporosis is greater the older the women; thus treatments solely related to fracture prevention are unlikely to be cost effective if targeted at pre or peri menopausal women. In previous evaluations of hormone replacement therapies for menopausal women the impact of the inclusion of fracture prevention has been minimal (e.g. Coyle et al. 2003d). For analysis specific to the Canadian health care system, treatment of osteoporosis prior to age 65 is unlikely to be covered by the drug formularies of the provincial ministries.

[44] Hormone replacement therapy (HRT) had been a commonly prescribed treatment for osteoporosis. However, given the continued evidence concerning the side effects associated with its use, HRT is no longer recommended solely for the treatment of osteoporosis and its use in this area has fallen dramatically.

water and there are similar restrictions relating to the time gap between taking therapy and eating.

In 2002, the Ontario Drug Benefit formulary allowed the prescribing of etidronate to all osteoporotic women covered by the provincial drug benefit program [45] (Ontario Ministry of Health and Long Term Care 2002). The use of alendronate and risedronate was restricted to women who had failed on previous treatments (i.e. etidronate)[46]. However, the Ontario Drug Benefit Formulary (ODB) had as an objective the revision of the prescribing guidance to more fully reflect the relative costs and benefits of each treatment option.

5.3 PREVIOUS GUIDELINES FOR ECONOMIC EVALUATIONS RELATED TO OSTEOPOROSIS

5.3.1 Introduction

Four separate reports providing guidance on the conduct of economic evaluations of treatments specific to osteoporosis have been identified.

5.3.2 Group for the Respect of Ethics and Excellence in Science

A 1997 report from the Group for the Respect of Ethics and Excellence in Science provided recommendations for evaluation of drugs registered for the prevention of treatment of osteoporosis (Dere et al. 1998). The group consisted of leading health economics and clinicians. However, details of the process by which

[45] The program provides drug benefits to seniors (65 years of age and older) and families on social assistance.

[46] Failure was characterized by either failure to maintain bone density after 1 year, failure to tolerate or incidence of fracture

recommendations were made were not provided. Major recommendations related to the need for modeling, the adoption of a long term time horizon, the need to include jurisdiction specific data on costs and epidemiology, inclusion of long term costs associated with fractures and the incorporation of quality of life effects.

5.3.3 WHO Collaborating Centre

The World Health Organization provided similar recommendations on the conduct of evaluation in osteoporosis (WHO Collaborating Centre 1999). The group consisted of leading clinicians with some overlap with the previous report. Details of the process of reaching consensus were not provided. Most of the recommendations are generic to all economic evaluations not just those relating to osteoporosis. Osteoporosis specific recommendations related to the inclusion of long term costs, the need to estimate age specific fracture rates, the need to consider any ongoing benefit of treatment after treatment is stopped and the need to base economic models on the reduction in fractures rather than intermediate outcomes such as bone mineral density.

5.3.4 Tosteson and Colleagues

Tosteson and colleagues identified eight methodological challenges in the conduct of economic analyses in osteoporosis (Tosteson et al. 2001a). The group comprised leading health economists and clinicians and was funded by both the National Institutes of Health and Proctor & Gamble Pharmaceuticals. The methods for reaching consensus were not provided.

Recommendations relating to each challenge were as follows: a preference for fracture based studies rather than bone mineral density (BMD) based modeling; the need to include health states incorporating all impacts of disease and treatment; the need to include mortality following fracture; the need to account for the costs of long term care; the need to include health utilities; the need to model any benefit beyond therapy; the need to validate a model by ensuring it is calibrated (i.e. replicates population incidence and prevalence rates); and the need to include data specific to the population under consideration.

5.3.5 OMERACT

5.3.5.1 Introduction

The final guidance document relates to the recommendation of a reference case for economic evaluations in osteoporosis which was developed as part of the OMERACT initiative (Coyle and Tosteson 2003)[47]. The OMERACT initiative is an international collaboration interested in defining appropriate outcome measurement in the area of rheumatology. A major focus of this initiative has been to establish reference cases for evaluations in different clinical areas (Maetzel et al. 2003). At the 2000 and 2002 conferences, progress was made towards developing a reference case for economic evaluations in osteoporosis. Issues relating to the conduct of economic evaluations were identified and recommendations on appropriate methods for dealing with these issues were obtained by debate between leading clinicians and health economists working in this area.

[47] I was the lead author on the report of the reference case and was involved in all stages of its development. The report reflects the consensus views of those participating in the process. Given the methods adopted in developing the guidance and my role in its development, recommendations from this report are provided in greater detail.

Recommendations were divided into two categories; those for which consensus emerged and those for which further debate was required. For the latter tentative recommendations were made.

5.3.5.2 Issues of Consensus

Issues which were identified as having a consensus recommendation were as follows:

Study Purpose and Population

Studies should report patient characteristics relevant to the evaluation of therapies for osteoporosis. Important characteristics include age and whether the patient had previous fracture; both strong predictors of the baseline risk of fracture (Papadimitropoulos et al.1997, Klotzbuecher et al. 2000). If possible, studies should incorporate a stratified analysis where the costs and benefits of therapies are estimated for alternate patient profiles (Coyle et al. 2003a).

Clinical Data

The effectiveness of therapies should be based on efficacy data from clinical trials. Where possible, evaluations should be conducted based on the results of meta analysis rather than single trials as this limits the potential for bias (Coyle and Lee 2002). When conducting such analyses, attention should be given to the use of bone mineral density t-scores for trial inclusion. Baseline rates of fractures and mortality should be obtained from relevant population databases for the geographical location for which the analysis is being conducted (Papadimitropoulos et al. 1997, Jacobsen et al. 1990, Johnell et al. 1992).

Resource Use

Evaluations should consider the costs of drug therapies including health care costs associated with monitoring of drug therapies (e.g., additional health care provider visits) and/or managing treatment-emergent side-effects. In addition, both the acute and long-term costs associated with fracture should be included in analyses. If any extraskeletal effects of the treatment are documented, then related resource use should be included in the analysis.

Discounting

Future costs and benefits should be discounted adopting a discount rate accepted by the jurisdiction to which the study is aimed: e.g. 0, 3 and 5% (CCOHTA 1997, Gold et al. 1996).

Source of Study Funding

There has been concern over the potential bias from studies funded by the pharmaceutical industry (Hillman et al. 1991, Kassirer and Angell 1994). However, it is unlikely that sufficient economic studies could be produced without industry funding. Therefore, in addition to a statement on funding source, authors of industry-sponsored studies must demonstrate their independence in the conduct and reporting of the economic evaluation. Only studies that were conducted under contracts that allowed for their independence over all aspects of study design, analysis, interpretation and reporting of results should be considered for publication.

5.3.5.3 Issues of Debate

The following issues were identified as ones which were still open to debate. For each, the tentative recommendation from the OMERACT report is provided:

Study Perspective

Where possible, studies should adopt the societal perspective. The impact of adopting a societal perspective will be dependent on whether the productivity losses associated with informal caregivers are deemed appropriate to include. Treatment of osteoporosis is ostensibly for patients who are past working age. Thus, if costs related to informal caregivers are excluded, evaluations incorporating costs to both the health care and social care sectors should be accepted as close approximations to the societal perspective (Goeree et al. 1996). However, if informal caregiver costs are included their incorporation may lead to lower cost effectiveness ratios than from a health care perspective.

Recommendation

Studies should at least adopt a perspective incorporating costs to the health and social care systems. Analysts should be encouraged to adopt the societal perspective and further studies should be conducted to estimate informal caregiver costs.

Basis of Modeling Osteoporosis Outcomes

Previous models used in conducting economic analysis in osteoporosis can be categorized as either age-specific fracture incidence based models or BMD based models (Tosteson et al. 2001a, Zethraeus et al. 2003b). To model treatment effectiveness, fracture incidence-based models directly apply the relative risk

reduction for therapy reported in clinical trials to baseline age-specific fracture incidence rates in the population of interest.

Relative risk reductions from randomized controlled trials usually are reported separately for vertebral and non-vertebral fracture sites. However, for different non-vertebral fractures the proportion of fractures attributable to osteoporosis varies substantively which infers that the relative risk reductions from therapy will vary by the location of non-vertebral fractures (Melton et al. 1997). In addition, there are substantive differences in the costs, mortality and quality of life effects associated with fractures (Coyle et al. 2001a). Thus, for economic analysis it is necessary to obtain relative risk reductions for specific fracture sites; vertebral, wrist, and hip.

BMD based models use epidemiological evidence to parameterize fracture incidence as a function of BMD and age. A caveat to this approach is that the evidence provided in epidemiological studies linking BMD changes to fracture risk often reflect cross-sectional population differences and may not be valid for interpreting the likely effect of longitudinal BMD differences observed in clinical trials.

Thus the BMD based approach is more complex than the fracture-incidence based approach and has a greater potential of error. However, they have been attractive in that previously trials of osteoporotic therapy tended to focus on detecting differences in BMD rather than a decline in fracture rates. As evidence of reduction in fractures has become required by regulators the need for BMD based models is less clear (Committee for Proprietary Medicinal Products 2001). Furthermore, for newer osteoporotic treatments such as bisphosphonates it is unclear that a BMD level

obtained through treatment will be associated with the same level of fracture risk if it occurred without treatment (Cummings et al. 2002) Ultimately, for economic modeling and clinical trials planning purposes it would be desirable to develop a comprehensive model that could accurately predict fracture on the basis of both BMD changes and markers of bone resorption and formation.

Recommendation: We propose that future economic analyses of interventions for osteoporosis follow previous recommendations by adopting fracture based models (Tosteson et al. 2001a, WHO Collaborating Centre 1999). In addition, we propose that relative risks should be obtained for at least the three primary fracture sites and should be based on symptomatic fractures.

Mortality Following Fracture

There is convincing evidence of mortality post hip fracture (Cooper et al. 1993, Cree et al. 2000). However there is less convincing evidence of a mortality effect associated with vertebral fracture (Cooper et al. 1993, Kado et al. 1999).

Recommendations: We propose that economic evaluations in osteoporosis should incorporate a mortality effect associated with hip fractures and where possible such data should be based on the specific geographical location for which the study is conducted. For mortality following vertebral fractures we recommend analysis can be conducted with or without such effects and reiterate that further clinical research is needed in this area.

Long Term Care Admission Post Hip Fracture

Hip fractures are associated with increased admission to long term care facilities though this will vary by country due to differences in the availability and funding of such care (Cree et al. 2000). However, it is unclear whether prevention of fractures will reduce admission to such facilities or merely delay it. A population-based study from Olmsted County, Maine, USA suggests that savings due to nursing home stays averted through hip fracture prevention are likely to be overly optimistic (Leibson et al. 2002).

Recommendation: We propose that studies should incorporate data on long term care admission specific to the geographical location for which the study is conducted. Furthermore, we propose that analysis should be conducted based on two assumptions: that all future long term care costs can be attributed to fracture; and that only the costs of LTC in the first year post hip fracture are assumed to be directly attributable to fracture. The latter can be seen as a more conservative assumption that will bias against effective therapies. Further research addressing fracture-attributable length of stay in long term care is needed.

Lack of Head to Head Trials

Within economic evaluation, the cost effectiveness of therapies is assessed relative to other available interventions. The choice of comparator therapy is a major determinant of the results of an analysis. Existing guidelines tend to differ modestly in their preferred choice of comparator – however, they tend to favour adoption of usual practice as at least one of the comparators.

A major limitation in the conduct of economic analyses however, is the lack of head to head trials comparing the therapies of interest. This problem exists primarily as a result of the requirement for placebo controlled trials with respect to the licensing of pharmaceuticals. Thus, if we wish to compare treatment options, it is necessary to estimate the relative effects of treatments through synthesis of placebo controlled trials. The Australian guidelines for pharmacoeconomics disallow any claim of superiority for a pharmaceutical based on synthesis of trials (Commonwealth Department of Human Services and Health 1995). Other guidelines tend to have less rigorous positions with respect to this issue allowing comparisons through carefully designed synthesis.

Recommendation: We propose that head to head comparisons can be made through careful synthesis of similarly designed trials and application of model-based economic evaluation techniques. However, we also recommend that pharmaceutical manufacturers be encouraged to conduct head to head trials.

Incorporating Extraskeletal Effects

In addition to their impact upon osteoporotic fractures, therapies may have extraskeletal effects. Depending on the therapies under consideration within an analysis, the importance of such effects will vary. Thus, the selection of health states to include in model-based economic evaluations of osteoporosis treatment is one that warrants careful consideration. Until recently, postmenopausal hormone replacement therapy, a treatment with widely recognized beneficial (e.g. reductions in menopausal symptoms and protection against colorectal cancer) and harmful (e.g., increases in breast cancer, thromboembolic events) extraskeletal effects, was a

mainstay for osteoporosis prevention and treatment. Thus, model-based analyses for postmenopausal hormone replacement therapy required explicit attention to health states related to these extraskeletal effects. Failure to include the full complement of health states likely to be affected by a therapeutic agent could produce misleading economic evidence. For example, a treatment that reduced hip fracture incidence by 90%, but increased breast cancer incidence by 50% may appear as a great success unless the harms were appropriately modeled. For other therapies, incorporation of extraskeletal effects will have minimal effect on analysis. For example, with bisphosphonates a possible adverse effect is an increased risk of gastrointestinal problems which can be alleviated by discontinuation of therapy or improved adherence. Such effects can be considered by accurate modeling of treatment discontinuation rates.

Recommendations: We propose that in future studies, analysts consider the impact of therapies on extraskeletal effects and incorporate these as necessary to accurately assess the incremental cost-effectiveness of alternative treatments. Such effects should be incorporated adopting methods consistent with those discussed with respect to fractures.

Benefit beyond Therapy

There is evidence that patients experience continued reductions in the risk of fracture after stopping therapy (e.g. Tonino et al. 2000). Previous studies have assumed that a patient will experience continued benefit in terms of fracture reduction over a time period equal to therapy duration; although, it is assumed that magnitude of benefit will decrease linearly over this period (Tosteson et al. 2001a, Zethraeus et al. 2003b).

However, evidence of continued benefit comes from studies where the follow up of patients has been for no more than two years post treatment curtailment (Tonino et al. 2000).

Recommendation: We propose that future studies should conduct multiple analyses based on assumptions relating to benefit to be obtained beyond therapy duration. As a minimum, analysis should be based on assuming no benefit beyond treatment and a decline in benefit in terms of fracture reduction over a time period equal to either therapy duration or a period up to two years.

Model Validation

Model validation should focus on calibration; that is that the model replicates all population estimates for each individual parameter (Kuntz and Weinstein 2001). This is necessary because in certain instances, the specific data required for modeling are unavailable though sufficient data are available for the interpolation of such parameters. For example, age-specific mortality (excluding mortality post fracture) may be unavailable, though age specific all cause mortality and age specific mortality post fracture is available. Thus, a model should be calibrated such that the combination of mortality rates post fracture and mortality rates without fracture replicates age specific all cause mortality. Failure to replicate models can lead to a major overestimation of the benefits of treatment.

Recommendation: We propose that for future studies, all models are fully calibrated.

Compliance with Therapy

Consideration of patient compliance raises four specific issues to consider; the measurement of compliance versus continuation; how to measure compliance; how soon do patients obtain benefits of treatment given non compliance; and compliance beyond the duration of clinical trials.

It is generally much easier to measure whether patients have obtained prescriptions for medication rather than whether they have taken medications correctly. This is the distinction between continuation and compliance. The difference between these concepts should be recognized within an economic evaluation where the analyst often models continuation for lack of complete data on compliance. When clinical trials results are reported on an intention-to-treat basis, it is noted that estimates of treatment efficacy are likely to already accommodate the impact of treatment non-compliance.

Estimating compliance levels is problematic. Two distinct forms of approach are available; prescription based records and patient based reports. Prescription based records typically involve the use of administrative databases from health care insurers – for instance the Ontario Drug Benefit (ODB) Program's database can be used to estimate compliance with therapies at 6 and 12 months. These can be seen as measures of patient's continuation with therapy though this need not infer compliance. Alternatively, therapy use can be estimated through such measures as pill counts and diaries though again these need not be accurate measures.

There is a potential for clinical trials evidence on continuation to be more optimistic than what is found in routine clinical practice. Therefore, clinical trials evidence should be augmented with real world patient compliance information whenever possible. For example, a recent paper on early osteoporosis treatment discontinuation among women initiating treatment with low bone mineral density showed self-reported rates of discontinuation of approximately 1 in 4 for postmenopausal hormone therapy and 1 in 5 for raloxifene and alendronate (Tosteson et al. 2003).

Thus, it is necessary in economic evaluation to be concerned over when the benefits of therapy are likely to commence. In clinical trials reporting fracture reductions by year of study, there is clear evidence that for many therapies treatment effect begins within 1 year of therapy (Harris et al. 1999).

Most clinical trials of therapies for osteoporosis are conducted over a short period of time: 2-3 years. As duration of therapy can be longer than trial duration it is necessary to model the effects of therapy beyond the period for which efficacy data are available. Generally, previous studies have adopted the same relative risks of fractures beyond the trial duration. This seems justified given that there is evidence of continued benefits from therapies up to 7 years duration (Tonino et al. 2000).

Recommendations: In all evaluations, some empiric measure of compliance should be used, though sensitivity analysis based on different rates of treatment continuation is required. We recommend that for economic evaluations in osteoporosis, that it is assumed that for individuals taking therapy for less than 1

year no treatment effect is obtained but for individuals taking therapy for at least 1 year the full treatment effect is obtained. Furthermore, we recommend that for therapy taken beyond the duration of trials that benefits are assumed to continue to the same extent.

Incorporation of Utilities

The principal impact of the sequelae of osteoporosis is the detriment in the quality of life of individuals who suffer a symptomatic fracture. For economic evaluation this is best incorporated by obtaining utility values for the specific fracture health states. Typical health states will relate to the fractures modeled within the analysis; hip fracture, wrist fracture and vertebral fracture as well as a "normal health" state relating to the absence of fracture (Cranney et al. 2001a). With respect to utility measurement there are two specific areas where there exists a lack of consensus; the duration of quality of life effects associated with fracture; and what should be the preferred approach for obtaining utility weights.

Previous studies have typically assumed that the quality of life effects of vertebral and wrist fractures are limited to the first year post fracture. However, for hip fractures the quality of life effects may be longer lasting. In addition, previous model-based economic evaluations have typically assumed that the "worst" post-fracture health state is that associated with hip fracture. However, recent evidence has challenged this assumption by findings of lower health utility among persons with hip and vertebral fracture relative to those with hip fracture alone (Tosteson et al. 2001b).

There are several methods of obtaining societal utility values for osteoporotic health states (Torrance and Feeny 1989, Tosteson and Hammond 2002). Direct utility elicitation methods, such as the standard gamble and time trade off, can be adopted by developing health state scenarios for osteoporotic health states and obtaining utility values for these scenarios from the general public. Alternatively, an indirect approach to health state valuation can be undertaken by having osteoporotic patients complete a standardized utility questionnaire, which has been linked through construction of a scoring algorithm to societal health state values (Brazier et al. 2002). One study that considered effect of health state utility values on the cost-effectiveness of an intervention that reduced hip fracture incidence by 50% suggested that the two approaches could result in qualitatively different results (Gabriel et al. 1999).

Recommendations: In economic evaluations in osteoporosis, two distinct quality of life weights should be adopted for all fractures: one relating to the first year post fracture and a second relating to long term effects (this may for some fractures be equivalent to normal health). Indirect elicitation has the advantage of providing societal health state values to the full range of outcome health states experienced by individuals with osteoporosis. It is recommended that analysts treat all utility values with caution and conduct appropriately detailed sensitivity analysis.

Accommodating Uncertainty

Previous recommendations for sensitivity analysis for studies in rheumatoid arthritis suggest that the minimum requirement should be for simple one way analysis of the major clinical, costs and quality of life parameters (Maetzel et al. 2003). There have

been considerable developments in the methods of analyzing uncertainty in economic analysis (e.g. Briggs et al. 2002, Felli and Hazen 1999). Given the wide range of uncertainty concerning many parameters within an osteoporosis based economic evaluation, more advanced techniques for sensitivity analysis should be explored. Monte Carlo simulation techniques can both identify those variables that have major impact in the results of analyses and provide a more accurate expression than simple deterministic analysis of the expected value of outcomes of interest (Doubilet et al. 1985, Thompson and Graham 1996). Such techniques are more easily conducted through the development of appropriate analytical software.

Recommendation: As a minimum economic analysis in osteoporosis should adopt simple univariate and multivariate sensitivity analysis. However analysts should be encouraged to adopt advanced methods for analyzing uncertainty with preference for the use of Monte Carlo simulation techniques.

5.3.6 Summary of Recommendations

Defining standards for economic evaluations in osteoporosis should improve the quality of future studies and facilitate comparisons between studies. This should ultimately allow more efficient health care provision in this disease area. Each of the guidelines described are steps towards defining such standards.

The guidelines are remarkably consistent in their recommendations for the conduct of studies. The only substantive difference between them is the extent to whether a

particular issue is covered within the guideline. There are no apparent issues in which the guidance from one report is inconsistent to guidance from another.

Table 5.1 provides a summary of the recommendations regarding each issue raised in the conduct of studies. Recommendations are based on the content of the previous guidelines. For some issues the recommendations represent the consensus of all the guidelines, whilst for others they represent the recommendations of the specific guidelines which covered this issue. In the following section, these recommendations are used to determine the quality of existing studies and as a template for the economic model described in Chapter 6.

5.4 REVIEW OF PREVIOUS ECONOMIC EVALUATIONS

5.4.1 Introduction

This section reports the findings of a review of economic evaluations of osteoporosis treatments for post menopausal women[48]. Studies were identified through an electronic search in MEDLINE, EconLit and Current Contents for economic evaluations in the area of osteoporosis using the keywords: cost-benefit analysis, cost, cost-analysis, osteoporosis, and fractures. The search was supplemented by searching the reference list of relevant studies and by a complementary internet search. In addition, studies were identified from previously published review articles (Cranney et al. 1999, Zethraeus et al. 2003b, Sculpher et al. 1999).

The following data were extracted from studies:

- Country of origin

[48] Studies focusing solely on hormone replacement therapy were excluded given the role for this therapy in the management of osteoporosis has diminished.

Table 5.1: Recommendations for Economic Evaluations of Osteoporosis Interventions

Methodological Issue	Recommendation
Time horizon	Model should incorporate long term effects: preferably adopting a lifetime horizon.
Study perspective	Perspective should be at least that of the health care system.
Modelling fractures	Models should use age-specific fracture incidence with treatment effects modelled using relative risk reductions derived from meta analysis of RCTs.
Data sources	Data should be specific to the population of interest.
Mortality following fractures	At minimum should incorporate attributable mortality following hip fracture.
Long term care admission post fractures	Consider two extremes; that all future long term care costs are attributable to fractures and that only first year costs are attributable.
Head to head comparisons	Can be achieved through careful synthesis of similarly designed trials.
Extraskeletal effects	All effects which will impact cost effectiveness should be included.
Benefit beyond therapy	Consider two extremes: that there is no benefit beyond therapy and that benefit is maintained.
Model validation	All models should be fully calibrated.
Compliance with therapy	Empiric measures of discontinuation should be used.
Utility values	Utility values should be used to facilitate cost utility analysis.
Accommodating variability	Analysis should be conducted for patients with different characteristics. Stratified analysis is encouraged.
Accommodating uncertainty	At a minimum, univariate and multivariate sensitivity analysis. Probabilistic analysis is preferred.
Study funding	Source of study funding should be stated as should investigator independence and any potential further conflict of interest.

Source: Synthesis of Coyle and Tosteson (2003), Dere et al. (1998), Tosteson et al. (2001) and WHO Collaborating Centre (1999)

- Treatment comparators
- Study design (type of model or clinical trial)
- Patient characteristics
- Form of analysis (cost minimization, cost effectiveness or cost utility)
- Outcome measure
- Summary of results and conclusions

In addition, each study was assessed in terms of how they measured up to the recommendations detailed in Table 5.1.

5.4.2 Review of Studies

Table 5.2 is a summary of all identified economic evaluations of therapies for osteoporosis. Fifteen studies were identified; 5 from the UK, 3 Canadian, 2 each from Sweden and Denmark, and 1 each from the USA, Italy and Norway. Eight studies were cost utility analyses whilst the other seven were cost effectiveness analysis with hip fracture (4), vertebral fractures (2) and risk ratios (1) as outcome measures. Studies were published over a ten year period from 1994 to 2003; with six studies published in the first five year period and nine in the latter period.

Of the 15 studies, 12 included a bisphosphonate as a comparator. Of these 9 included alendronate, 2 included risedronate and 6 included etidronate. The cocnlusiosn of the studies including bisphosphonates were not consistent. Alendronate appears to be cost effective compared to etidronate in two studies (Coyle et al. 2001, Best et al. 1998). However, in two other studies the opposite conclusions were reached (Aursnes et al. 2000, Visentin et al. 1997). In the only

Table 5.2: Summary of Economic Evaluations for Treatments of Postmenopausal Women with Osteoporosis

Study/Country of Origin	Comparators	Design	Base patient characteristics	Form of analysis	Outcome	Summary of Results
Ankjaer Jensen (1996) Denmark	Calcitonin Calcium Etidronate HRT	Decision model	70 year old women screened and treated for osteoporosis	CEA	Hip fractures	Both HRT and etidronate cost saving and effective Calcium: ICER 7 609 DKK compared to no therapy Calcitonin: ICER 142 300 DKK compared to no therapy
Aursnes (2000) Norway	Alendronate Etidronate	Decision model	Osteoporotic women (age not given)	CEA	Risk ratios	ICERs nor presented: for alendronate to be as cost effective as etidronate needs to be priced 40-70% higher
Bendich (1999) USA	Calcium	Decision model	Women aged > 50	CEA	Hip fractures	Calcium supplementation will be cost saving and will reduce hip fractures
Best (1998) UK	Alendronate Etidronate	Decision model	Osteoporotic women (age not given)	CUA	QALYs	Alendronate: ICUR £40 500 Etidronate: ICUR £48 140
Coyle (2001a) Canada[49]	Alendronate Etidronate Nasal calcitonin	Markov model	Osteoporotic women aged 65 with previous fracture	CUA	QALYs	ICURs Nasal calcitonin vs. no therapy: $46 500 Nasal calcitonin vs. Etidronate: $32 600 Nasal calcitonin vs. alendronate: uncertain (either could be dominant)

[49] This evaluation used on an earlier version of the model described in chapter 6 and used in chapters 7 and 8.

Table 5.2: Summary of Economic Evaluations for Treatments of Postmenopausal Women with Osteoporosis (cont.)

Study/Country of Origin	Comparators	Design	Base patient characteristics	Form of analysis	Outcome	Summary of Results
Francis (1995) UK	Calcitonin Etidronate HRT	Decision model	Osteoporotic women (age not given)	CEA	Vertebral fractures	HRT: ICER £277 Etidronate: ICER £1 880 Calcitonin: ICER £25 013
Grima (2002) Canada	Alendronate Risedronate	Markov model	Osteoporotic women aged 65	CUA	QALYs	ICURs Risedronate vs. no therapy: $16 158 Risedronate vs. alendronate: dominant
Iglesias (2002) UK	Risedronate	Markov model	Osteoporotic women aged 75 with vertebral fracture	CUA	QALYs	Risedronate is dominant over no therapy
Johnell (2003) Sweden	Alendronate	Markov model	Osteoporotic women aged 71 with previous vertebral fracture	CUA	QALYs	ICUR: alendronate SEK 76 000
Jönsson (2002) Denmark	Alendronate	Markov model	Osteoporotic women aged 71 with previous vertebral fracture	CUA	QALYs	ICUR: alendronate DKK 52 311
Kanis (2002) UK	Alendronate Calcitonin Calcium Fluoride HRT Raloxifene Vitamin D	Markov model	Osteoporotic women aged 50, 60, 70 and 80	CUA	QALYs	ICURs compared to no therapy for 70 year old woman: Alendronate £13 300 Vitamin D £22 700 Calcitonin £168 000 Calcium £15 900 Fluoride - dominated by no therapy HRT £24 100 Raloxifene £187 000

Table 5.2: Summary of Economic Evaluations for Treatments of Postmenopausal Women with Osteoporosis (cont.)

Study/Country of Origin	Comparators	Design	Base patient characteristics	Form of analysis	Outcome	Summary of Results
Rosner (1998) Canada	Strategies incorporating Alendronate Calcium Etidronate HRT	Markov model	Postmenopausal osteoporotic women	CEA	Vertebral fractures	If unwilling to spend more than $166 to prevent vertebral fracture a strategy of calcium only is optimal. If willing to spend between $166 and $2 331 a strategy of HRT then calcium is optimal. If willing to spend between $2331 and $40 965 a strategy of HRT then etidronate then calcium is optimal. If willing to spend more than $40 965 a strategy of HRT then alendronate then calcium is optimal.
Torgerson (1995) UK	Vitamin D Calcium	Decision Model	Elderly osteoporotic women	CEA	Hip fracture	ICER Parenteral Vitamin D: £2 317 Oral vitamin D + calcium: £22 379
Visentin (1997) Italy	Alendronate Calcitonin Clodronate Etidronate Fluoride HRT Vitamin D	Decision model	Osteoporotic women (age not given)	CEA	Hip fracture	ICER Alendronate: US$350 628 Calcitonin: US$574 900 - $1 044 416 Clodronate: US$804 142 Etidronate: US$141 082 Fluoride: US$281 147 HRT: US$51 456 - $81 416 Vitamin D: US$105 812 - $110 948
Willis (2002) Sweden	Calcium + Vitamin D	Markov model	Osteoporotic women aged 70	CUA	QALYs	Calcium + Vitamin D cost saving and beneficial

study comparing risedronate to another bisphosphonate (alendronate); risedronate was cost effective.

When drawing conclusions from the published studies it is necessary to consider that the overall quality of studies when assessed against the 15 recommendations for good practice from the previous section was poor (Table 5.3). The highest quality study was the recent UK NHS HTA report on osteoporotic treatments (study #11 in Table 5.3) where 13 of the 15 desired standard were met (Kanis et al. 2002). However, the study failed to both explicitly provide head to head comparisons between therapies and to model discontinuation rates based on actual data[50].

In other studies, the number of standards met ranged from 1 to 11. Only one study, adequately allowed for drug discontinuation (Rosner et al. 1998). Other recommendations which were rarely met related to adequate modeling for fractures (only 3 studies), modeling head to head comparisons (3) and incorporating benefit beyond therapy (3). Given the lack of quality of studies, it is difficult to assume that they provide any substantive evidence towards assessing the cost effectiveness of bisphosphonates for the Canadian situation.

Three studies did not include a bisphosphonate as a comparator focusing on calcium with or without vitamin D (Bendich et al. 1999, Torgerson et al. 1995, Willis et al. 2002). In all studies calcium was found to be both effective and cost saving. Of the

[50] In the analysis detailed in Chapters 7 and 8 all 15 recommendations for the conduct of studies are met and in addition analysis incorporates a full consideration of uncertainty and variability through adoption of the normative framework as detailed in Chapters 3 and 4.

Table 5.3: Quality of Published Economic Evaluations of Osteoporosis Interventions

	Study Number														
Methodological Issue	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Time horizon	X	X	X	X	√	X	X	√	√	√	√	X	X	X	√
Study perspective	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
Modelling fractures	X	X	√	X	√	X	X	X	X	X	√	X	X	X	X
Data sources	√	X	√	√	√	X	X	X	√	X	√	X	√	√	√
Mortality following fractures	X	X	X	√	√	X	X	X	X	√	√	X	X	X	√
Long term care admission post fractures	X	X	√	X	X	X	X	X	X	√	√	X	X	X	√
Head to head comparisons	X	X	n/a	X	√	X	√	n/a	n/a	n/a	X	√	X	X	n/a
Extraskeletal effects	√*	n/a	n/a	n/a	√	X	n/a	n/a	n/a	n/a	√*	X	n/a	X	n/a
Benefit beyond therapy	X	X	X	X	X	X	X	X	√	√	√	X	X	X	X
Model validation	X	X	√	X	X	X	√	X	√	√	√	X	X	X	√
Compliance with therapy	X	X	X	X	X	X	X	X	X	X	X	√	X	X	X
Utility values	X	X	X	√	√	X	√	√	√	√	√	√	X	X	√
Accommodating variability	X	X	X	√	√	X	√	X	√	√	√	√	√	X	√
Accommodating uncertainty	√	X	X	√	√	X	√	X	√	√	√	√	X	X	√
Study funding	X	X	√	√	√	√	√	√	√	X	√	√	√	X	√

*** Analysis assumes beneficial effect of HRT in prevention of CHD**

1 Ankjaer Jensen (1996) 2 Aursnes (2000) 3 Bendich (1999) 4 Best (1998) 5 Coyle (2001a)
6 Francis (1995) 7 Grima (2002) 8 Iglesias (2002) 9 Johnell (2003) 10 Jönsson (2002)
11 Kanis (2002) 12 Rosner (1998) 13 Torgerson (1995) 14 Visentin (1997) 15 Willis (2002)

12 studies assessing the cost effectiveness of bisphosphonates, 9 included alendronate, 6 included etidronate and 2 included risedronate.

Both risedronate studies found the drug was cost effective. In one Canadian study, risedronate was dominant over alendronate and had an acceptable ICUR when compared to no therapy (Grima et al. 2002). In a UK study, risedronate was found to be dominant over no therapy (Iglesias et al. 2002). However, both studies had major problems relating to the methods for modeling fractures with and without therapy.

Five studies compared alendronate to etidronate. Two found etidronate to be more cost effective (Rosner et al. 1998, Visentin et al. 1997) whilst two other studies found alendronate to be more cost effective (Coyle et al. 2001a, Best et al. 1998). It was difficult to derive any meaningful interpretation from the one remaining study (Aursnes 2000).

5.5 CONCLUSION

In this chapter previous guidance for conducting economic analysis in osteoporosis were identified and summarized. The guidance documents were remarkably consistent in their coverage leading to fifteen recommendations concerning good practice. These recommendations relate to issues generic to all economic evaluations such as the lack of head to head studies, statements relating to study funding and analysis of uncertainty and variability. In addition, the recommendations include issues specific to economic evaluations of osteoporosis such as the modelling of fractures rather than bone mineral density and in the inclusion of mortality and long term care admissions post fracture.

A review of published economic evaluations identified 15 studies relating to the management of osteoporosis in postmenopausal women. Twelve of these studies included a bisphosphonate as a comparator. However, the paucity of good quality studies concerning the cost effectiveness of bisphosphonates in the management of osteoporotic women highlighted the need to conduct a full economic evaluation from the Canadian context. In the following chapter, a model designed to conduct such an evaluation is outlined. The model facilitates analysis which adheres to the recommendations for good practice contained in Table 5.1. In Chapters 7 and 8 the results of the evaluation are presented adhering to the normative framework for handling uncertainty and variability outlined in Chapters 3 and 4.

Chapter 6.

Economic Model of Osteoporosis in Canada

6.1 INTRODUCTION

Given the assumed benefit of bisphosphonates in reducing fractures, there was a need to assess the cost-effectiveness of treating osteoporotic women with bisphosphonates in the Canadian context. In Chapter 5, existing guidance for conducting economic analysis of osteoporotic therapies was reviewed. In addition, a review of published economic analyses found that previous studies had many methodological failings.

The objective of the latter half of this thesis is to conduct a full economic analysis of treatment options for osteoporosis adopting the normative framework outlined in Chapters 3 and 4. In this chapter an economic model of osteoporosis in Canada developed for this thesis is described. The model adheres to the recommendations described in the previous chapter.

6.2 MODEL STRUCTURE

A decision analytic model for osteoporosis was developed for this thesis[51]. The model reflects the natural history of women with osteoporosis incorporating the

[51] Previous versions of the model have been used with respect to number of published papers for which in all cases I have been the principal author for the economic analyses (Coyle and Lee 2002, Coyle et al. 2001a, Waldegger et al. 2003). The Coyle et al. (2001a) paper is reviewed as part of chapter 4. The Coyle and Lee (2002) paper uses the analysis from the earlier paper to highlight issues relating to the conduct of economic analysis. The Walldeger et al. (2003) paper uses the original model to determine the cost effectiveness of hip protectors specific to the elderly population within long term care.

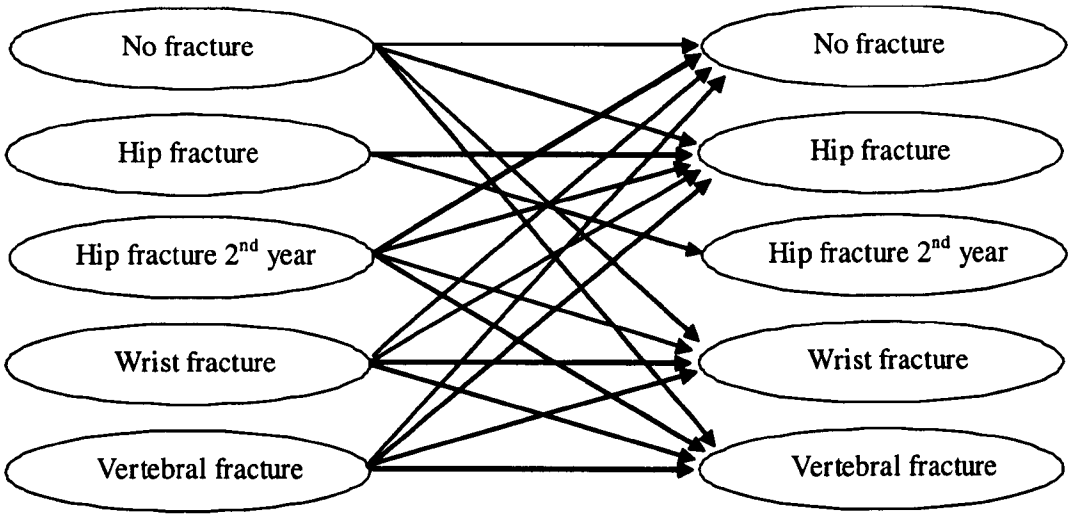
sequelae associated with osteoporosis (e.g. fracture) and also the transition of women both in terms of the development of osteoporosis, history of fracture and residential status (Figure 6.1). The model uses the most recently available data relevant to the Canadian population. Given the chronic nature of osteoporosis the model is a Markov model with a 1 year cycle length.

The probability of a woman experiencing a hip, wrist or vertebral fracture is assumed to be dependent on three factors; age, osteoporotic status and previous history of osteoporotic fractures (Papadimitropoulos et al. 1997, Marshall et al. 1996, Klotzbuecher et al. 2000). Both hip and vertebral fractures are associated with excess mortality and hip fractures are also associated with increased admission to long term care facilities (LTC) (Papadimitropoulos et al. 1997, Kado et al. 1999, Cree et al. 2000). In addition, the probability of hip fracture and the probability of mortality post hip fracture increases if a women resides within LTC (Papadimitropoulos 2000).

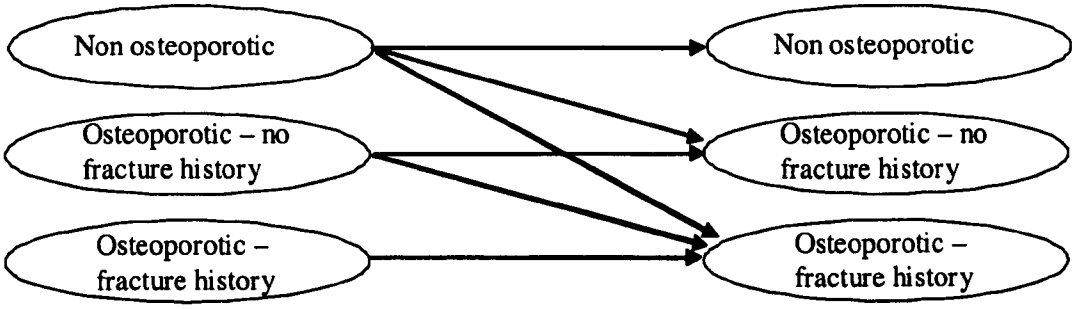
Thus, within the model it is necessary to distinguish between women on the basis of whether they have a history of fracture and whether they are living in the community or LTC. Based on these dichotomies; there will be different age specific transition probabilities for fracture (hip, wrist and spine) and associated risks of mortality from fracture (hip and spine only).

The model is populated with relevant transition probabilities and estimates of the costs and utilities associated with each health state (Tables 6.1, 6.2 and 6.3). Input parameters estimated through sample information rather than population estimates are represented by a probability density function based on their expected value and

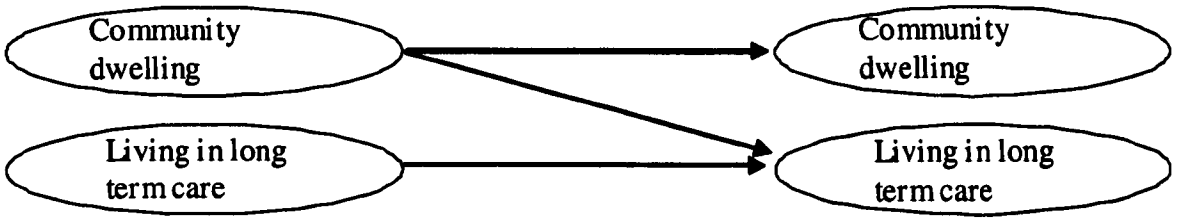
Figure 6.1: Design of Decision Model



a. Fractures



b. Osteoporotic status



c. Residential Status

Table 6.1: Parameter Estimates: Population Based Data

Parameter		Estimate
Annual probability of death : all cause		Age specific
Probability of mortality post hip fracture: community		$0.010*e^{0.084*age}$
Probability of mortality post hip fracture: long term care		$0.055*e^{0.072*age}$
Annual probability of hip fracture: community		$0.0578*e^{0.096*age}$
Annual probability of hip fracture: long term care		$73.99*e^{0.049*age}$
Annual probability of vertebral fracture	50-59	0.000543
	60-69	0.000899
	70-79	0.001920
	>80	0.004383
Annual probability of wrist fracture	50-59	0.000814
	60-69	0.003152
	70-79	0.004021
	>80	0.004380
Proportion of vertebral fractures requiring hospitalization	60-69	0.30
	70-79	0.38
	>80	0.45
Proportion of women residing in LTC	65-69	0.68%
	70-74	1.57%
	75-79	4.19%
	80-84	10.71%
	>85	29.93%
Proportion of women who are osteoporotic	50-59	0.06
	60-69	0.182
	70-79	0.27
	>80	0.421
Utility value for no current fracture	65-69	0.86
	70-74	0.83
	75-79	0.79
	80-84	0.70
	>85	0.55

Source: Sources provided in text.

Table 6.2: Parameter Estimates: Sample Data

Parameter		Estimate	Probability Density Function
Probability of admission to LTC post hip fracture	65-74	0.056	Beta (5, 85)
	75-84	0.166	Beta (26, 131)
	>85	0.298	Beta (25, 59)
Relative risk of fracture for reduction in BMD equivalent to one standard deviation of the young adult mean		1.5	Lognormal (1.5, 1.03)
Relative increase in risk of fracture for osteoporotic women with previous fracture		1.32	Lognormal (1.32, 1.14)
Relative risk of mortality post vertebral fractures		1.16	Lognormal (1.16, 1.06)
Attributability of long term care post fracture to fracture		0.5	Beta (1, 1)
Cost of hip fracture			
Women living in the community	65-74	22 124	Normal (22 124, 2 936)
	75-84	27 801	Normal (27 801, 2 562)
	>85	27 301	Normal (27 301, 3 241)
Women moving to LTC	65-74	43 459	Normal (43 459, 13 875)
	75-84	49 136	Normal (49 136, 3 899)
	>85	48 636	Normal (48 636, 4 122)
Women who remain living in LTC	65-74	15 699	Normal (15 699, 8 497)
	75-84	21 376	Normal (21 376, 5 374)
	>85	20 876	Normal (20 876, 4 184)
Women who die following fracture		15 498	Normal (15 498, 1 432)
Cost of wrist fracture		275	Normal (275, 69)
Cost of vertebral fracture			
ambulatory		128	Normal (128, 32)
hospitalized		4646	Normal (4 646, 1 162)
Utility values			
hip fracture		0.536	Normal (0.536, 0.037)
wrist fracture		0.976	Normal (0.976, 0.022)
vertebral fracture		0.674	Normal (0.674, 0.046)

Source: Sources provided in text.
Note: Beta distributions depicted by number of events and non events; normal and lognormal distributions by mean and standard error of the mean.

Table 6.3: Parameter Estimates: Treatment Specific Data

Parameter	Estimate	Probability Density Function
Annual cost of drug therapy		
alendronate	$730.59	Fixed
etidronate	$189.72	Fixed
risedronate	$692.49	Fixed
Continuation rates with therapy		
alendronate	0.65	Beta (65, 35)
etidronate	0.57	Beta (57, 43)
risedronate	0.62	Beta (62, 38)
Coefficient for benefit beyond therapy curtailment (0= linear decline)	0	Normal (0, 1)
Relative reduction in hip fractures		
alendronate	0.465	Lognormal (0.465, 1.29)
etidronate	0.945	Lognormal (0.945, 2.32)
risedronate	0.739	Lognormal (0.739, 1.13)
Relative reduction in wrist fractures		
alendronate	0.551	Lognormal (0.551, 1.27)
etidronate	0.818	Lognormal (0.818, 1.74)
risedronate	0.658	Lognormal (0.658, 1.27)
Relative reduction in vertebral fractures		
alendronate	0.439	Lognormal (0.439, 1.16)
etidronate	0.443	Lognormal (0.443, 1.18)
risedronate	0.665	Lognormal (0.665, 1.11)

Source: Sources provided in text and Table 6.4.
Note: Beta distributions depicted by number of events and non events; normal and lognormal distributions by mean and standard error of the mean.

the associated uncertainty. The probability density functions represent the likelihood of alternative population estimates for the parameters of interest.

As with the case study in Chapter 4, both drug prices and input values obtained from population rather than sample data are assumed fixed[52]. The model is developed within a Microsoft Excel 2000 spreadsheet incorporating the Crystal Ball software enhancement to facilitate Monte Carlo simulation (Crystal Ball 2000). The Monte Carlo simulation (MCS) involves obtaining several outcome estimates by re-running the model employing different values for each data input randomly selected from that variable's probability density function (Doubilet et al. 1985).

6.3 TRANSITION PROBABILITIES

6.3.1 Introduction

Given the structure of the model outlined above, the following transition probabilities are required to allow a simulation of progression through the model:

- Probability of developing osteoporosis specific to a woman's age
- Probability of being admitted to LTC specific to a woman's age
- Probability of hip, wrist and spine fracture specific to a woman's age, residential status, osteoporotic history and treatment.
- Probability of mortality with and without fracture specific to a woman's age

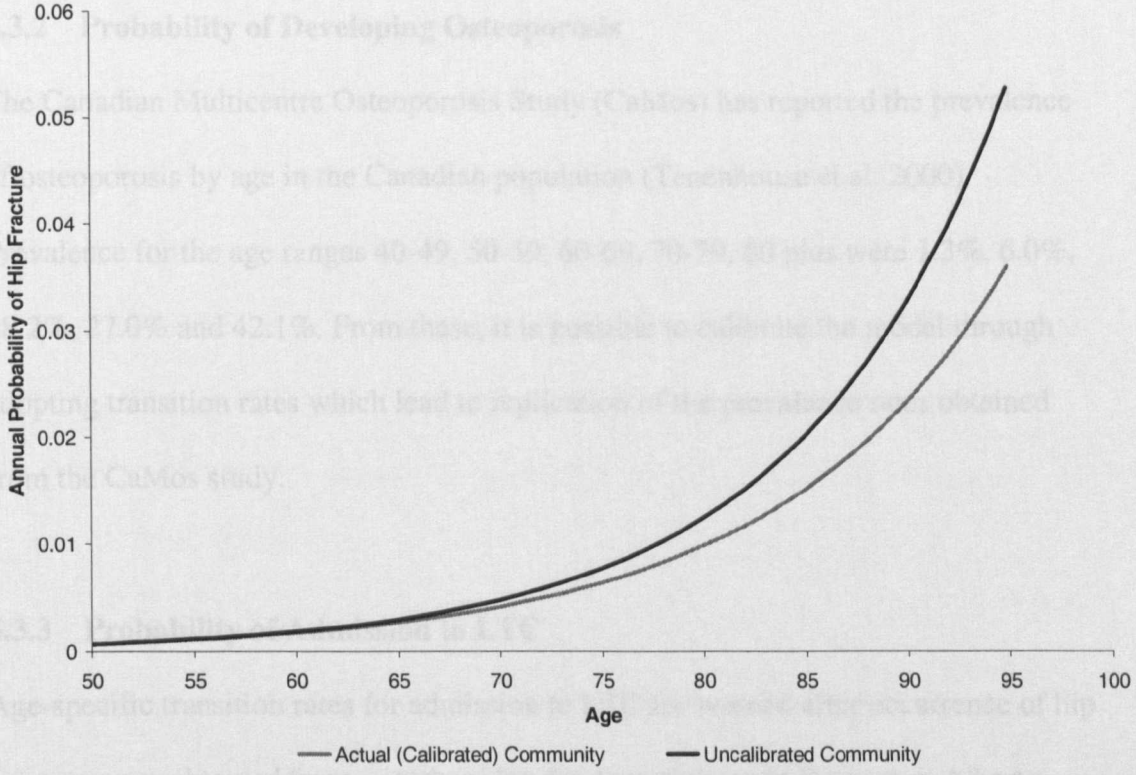
[52] Drug prices are assumed fixed as they represent the true cost of drug acquisition at the time of analysis. Population data are assumed fixed for two reasons. First, as they are based on population data it is unclear what the value of finding out the need for further information when such data cannot be obtained. Secondly, when data are based on population data any uncertainty around such inputs would be minimal.

In certain instances, the required data are unavailable. For example, in the model age specific rates of hip fractures for community dwelling woman are required for three classifications of women: non osteoporotic women, osteoporotic women without previous osteoporotic fracture and women with previous osteoporotic fracture. However, alternative data parameters are available which allow computation of the necessary parameters through calibration of the model (Kuntz and Weinstein 2001). The underlying age specific rate of hip fracture in the community is known. In addition, the relative risk of fracture for women who are osteoporotic compared to those who are not and the relative risk of fracture for osteoporotic women given a previous osteoporotic fracture versus no previous fracture are known. If the model merely involved weighting the underlying rate by the two relative risks this would lead to a major overestimate of the rate of fracture in the community (Figure 6.2).

Over-estimation would occur for two reasons. First, given that a large proportion of the fractures in the community occur in osteoporotic women with or without previous fracture the rate for women with no such risk factors will be substantially lower than the estimate from the whole population. Secondly, a proportion of osteoporotic women will have previous fracture; thus the relative risk of having a fracture when osteoporotic without fracture will be lower than the estimate available which compares osteoporotic women with and without fracture to non osteoporotic women.

Calibration involves adjusting the estimates of data such that the population data available is reproduced. In this instance failure to adjust the underlying rates of fracture would bias results in favour of active treatment as the underlying rates of

Figure 6.2: Effect of Failure to Calibrate Model on the Annual Probability of Hip Fracture for Community Dwelling Women



fractures would be over-estimated which in turn would lead to an over-estimate in the number of fractures avoided and the subsequent costs and outcomes[53].

6.3.2 Probability of Developing Osteoporosis

The Canadian Multicentre Osteoporosis Study (CaMos) has reported the prevalence of osteoporosis by age in the Canadian population (Tenenhouse et al. 2000). Prevalence for the age ranges 40-49, 50-59, 60-69, 70-79, 80 plus were 1.3%, 6.0%, 18.2%, 27.0% and 42.1%. From these, it is possible to calibrate the model through adopting transition rates which lead to replication of the prevalence rates obtained from the CaMos study.

6.3.3 Probability of Admission to LTC

Age-specific transition rates for admission to LTC for women after occurrence of hip fracture were obtained from a study of hip fracture patients in Edmonton Alberta (Cree et al. 2000). Analysis was based on 338 patients who lived in the community prior to fracture and who survived the immediate period post fracture. The rates of admission to LTC following fracture for patients aged 65-74, 75-84 and 85 plus were 5.6% (n=90), 16.6% (n=157) and 29.8% (n=184) respectively. Probability density functions for these variables assume a beta distribution based on the number of patients within each age cohort.

Estimates of the probability of residing in LTC (i.e. prevalence) are available for fiscal year 1993/94 based on data from Statistics Canada (Papadimitropoulos 2000). For age groups 65-69, 70-74, 75-79, 80-84 and 85+ the probabilities of residing in

[53] The impact of failure to calibrate on the results of this analysis will be presented at the 2004 meeting of the Society of Medical Decision Making (Appendix C).

LTC were 0.68%, 1.57%, 4.19%, 10.71% and 29.93%. Thus, the model is calibrated by adopting transition rates which allow replication of the probabilities of residing in LTC.

6.3.4 Probability of Fracture

The probability of hip fracture for women both living in the community and in LTC were derived from data for 1993/94 from both the Canadian Institute of Health Information and Statistics Canada (Papadimitropoulos et al. 1997, Papadimitropoulos 2000). The age specific risk of fracture per 100 000 community dwelling women was $0.578 * e^{0.096*age}$. For women living in LTC the age specific risk was higher : $73.99 * e^{0.049*age}$.[54,55]

Vertebral fracture rates were based on data from the Manitoba Health Services Insurance Plan for the years 1981 to 1984 (Hu et al. 1996). The study used data to derive age specific estimates for the rate of vertebral fracture. The study incorporated both ambulatory and hospital care of vertebral fractures. The annual rates of vertebral fracture per 100 000 women aged 50-59, 60-69, 70-79 and 80 plus were 54.3, 89.9, 192, and 438.3 respectively.

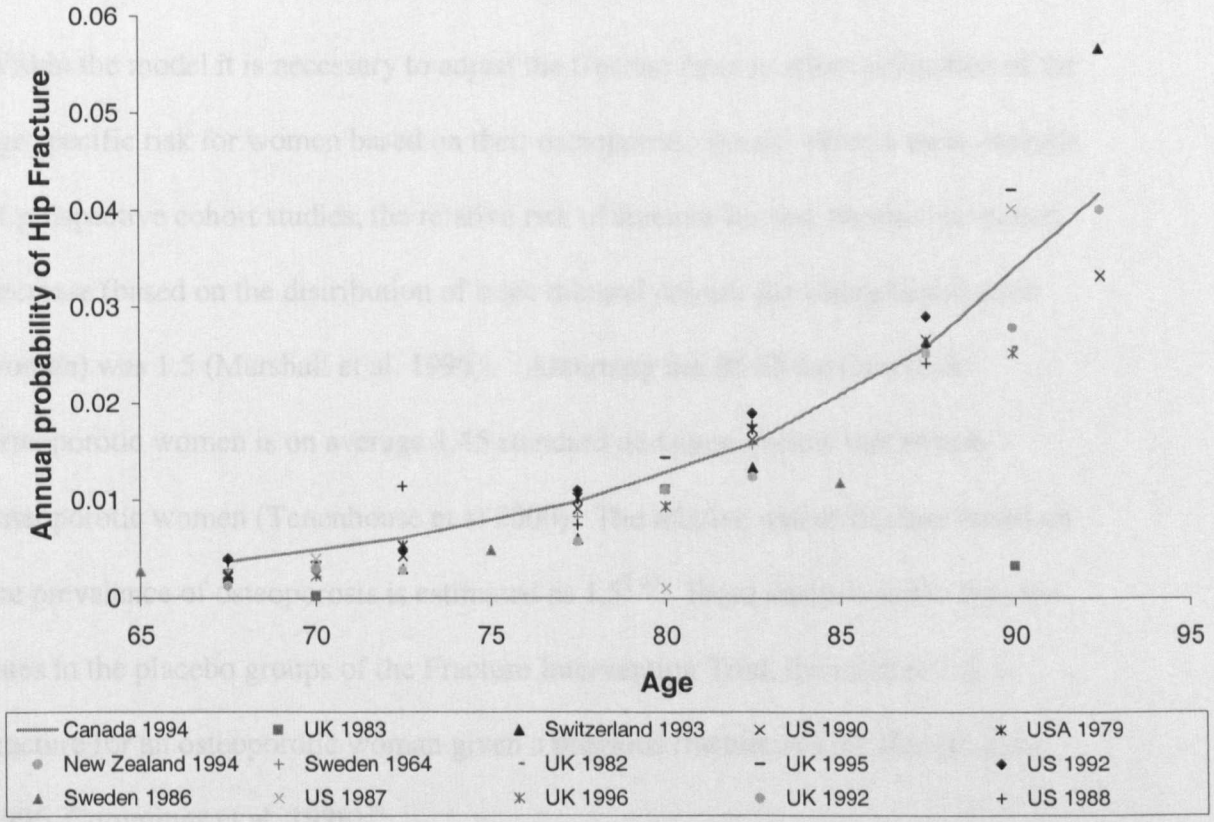
The probability of wrist fractures in the total population was derived from the same sources as used for vertebral fractures but for the dates 1986-1990 (Coyle et al.

---

[54] Given the paucity of data on women aged over 95, the risk of fracture was assumed to be constant above 95.

[55] Other studies have found a similar exponential increase in hip fractures with age although this is the only published study for Canada (Figure 6.3). The Canadian figures are consistent with other studies. Estimated rates of hip fracture before aged 75 are slightly higher than the median but after age 75, estimates are closer to the median reported rates.

Figure 6.3: Rates of Hip Fracture by Age, Country and Final Date of Population Sample



2001a). The annual rates of wrist fracture per 100 000 women aged 50-59, 60-69, 70-79 and 80 plus were 81.4, 315.2, 402.1 and 438.0 respectively.

Within the model it is necessary to adjust the fracture rates to allow estimation of the age specific risk for women based on their osteoporotic status. From a meta analysis of prospective cohort studies, the relative risk of fracture for one standard deviation decrease (based on the distribution of bone mineral density for young health adult women) was 1.5 (Marshall et al. 1996). Assuming the BMD for Canadian osteoporotic women is on average 1.45 standard deviations below that of non-osteoporotic women (Tenenhouse et al 2000). The relative risk of fracture based on the prevalence of osteoporosis is estimated as $1.5^{1.45}$. From analysis of the fracture rates in the placebo groups of the Fracture Intervention Trial, the relative risk of fracture for an osteoporotic woman given a previous fracture is 1.32 (Black et al. 1996, Cummings et al. 1998)[56].

Simply weighting the age-specific probability of fracture by these relative risks will lead to an over-estimate of the rate of hip fracture in each residential status category. Similarly, weighting the probability of hip fracture for osteoporotic women without fracture will be lower than the estimate from the Marshall study as a high proportion of women with osteoporosis have previous fracture. Thus, rates for the baseline risk of fracture in the population are imputed which allow replication of the age-specific probability of fracture in the population as a whole as well as allowing replication of relative risks associated with previous fracture history and osteoporosis.

---

[56] The Fracture Intervention Trial was a major clinical trial of the impact of alendronate on osteoporotic fracture. The trial was stratified by previous fracture history and comparison of fracture rates in the trial were used to determine the relative risk of fracture with previous fracture.

6.3.5 Mortality Rates

Mortality rates immediately post hip fracture for Canadian women were obtained from the same study as the probability of hip fracture (Papadimitropoulos et al. 1997, Papadimitropoulos 2000). The age specific mortality rates per 100 000 community dwelling women following fracture were $0.010 * e^{0.084*age}$. For women living in LTC, the relevant formula is $0.056 * e^{0.072*age}$.[57]

Canadian mortality rates immediately post vertebral fracture are unavailable. The results of a US prospective cohort study of 9704 women aged over 65 were used within the model (Kado et al. 1999). Within this study, the relative risk of mortality one year post vertebral fracture was 1.16.

Age specific all cause mortality rates for women are available for Canada for 1990-1992 (Statistics Canada 1995a). The model was calibrated by adopting age specific mortality rates for women without fracture which allow replication of the age specific all cause mortality rates for the whole Canadian woman population, given the effect of hip and vertebral fracture on mortality and their respective incidences.

6.4 COSTS

6.4.1 Introduction

Costs for each particular state within the decision model were estimated. The model adopts the perspective of the health and social care system in that the costs of health, social services and long term care are included. The analysis will approximate

[57] Given the paucity of data on women aged over 95, the risk of death following hip fracture was assumed to be constant after 95.

results from a societal perspective in that the productivity losses associated with informal care post fracture have been shown to be minimal (Goeree et al. 1996)

6.4.2 Hip Fractures

The first year cost of the health care resources associated with the treatment of a hip fracture comprises immediate acute care, rehabilitation and institutionalization. Costs associated with hip fracture beyond one year are discussed in Section 6.4.5.

Estimates were based on a study of the costs of hip fracture for 504 patients in Hamilton, Ontario (Wiktorowicz et al. 2001). Costs for individual patients were estimated through a review of patient hospital records linked with data from community services and long term care facilities. The coefficients from a multivariate analysis are used to estimate the mean costs for each sub-group for survivors post hip fracture. The costs for patients for whom death is attributable to fracture is provided directly and is assumed to be the same for all age groups ($15 498).

The probability density function for the cost of hip fractures is assumed to take a normal distribution with the standard error of the mean for each sub-group based on the standard deviation for all survivors adjusted by the sample within each sub-group. For deaths post hip fracture the function is based on the actual mean and standard error for this group.

6.4.3 Vertebral Fractures

For women with vertebral fracture, the proportion requiring hospitalization varies by age: 40%, 30%, 38% and 45% for ages 50-59, 60-69, 70-79 and 80 plus respectively (Hu et al. 1996).

The costs per inpatient stay for women with vertebral fractures treated on an inpatient basis were obtained from analysis of data from the Ontario Case Costing project for one teaching hospital and seven community hospitals (Ontario Case Costing Initiative 2000). The mean cost per hospitalized fracture patient was $4646 inflated to 2001 Canadian dollars. The cost of treating vertebral fractures on an ambulatory basis is estimated to be $128 based on the estimated resource use obtained from a survey of Canadian physicians (Ontario Ministry of Health and Long Term Care 2001, Goeree et al. 1996).

Following Briggs and colleagues (2002), when the degree of uncertainty is unknown, the probability density functions for costs are assumed to form a normal distribution with a standard error equivalent to a fixed percentage of the mean value. For the costs of a vertebral fracture treated in hospital or treated as an outpatient the standard error was assumed to be equivalent to 25% of the expected value.

6.4.4 Wrist Fractures

Wrist fractures are assumed to be exclusively treated on an ambulatory basis. The cost of treating wrist fractures was estimated to be $275 calculated on the same basis as vertebral fractures (Ontario Ministry of Health and Long Term Care 2001, Goeree et al. 1996). As with the costs of vertebral fractures, the probability density function

was assumed to be a normal distribution with a standard error of the mean equivalent to 25% of the expected value.

6.4.5 Long Term Care

The annual cost of long term care is assumed to be the difference between the annual costs following hip fracture for those staying in the community and those being admitted to long term care ($21 335). The model only incorporates the incremental costs of long term care occurring post hip fracture. It should be noted that although a patient may be admitted to long term care not all the stay may be attributable to hip fractures as the patient may have been admitted over time for other causes. Thus, the proportion of long term care stay attributable to fracture is assumed unknown with a mean of 0.5 (Beta 1,1) where 0 means only the first year post fracture is attributable and 1 means all of the subsequent long term care stay is attributable[58]. The two extremes of this distribution are equivalent to the scenarios recommended within the OMERACT reference case discussed in the previous chapter (Coyle and Tosteson 2003). Thus, the model adopts an assumption that the true proportion is uncertain and lies between these extremes.

6.5 UTILITIES

Utility values were required for the following health states: normal health, hip fracture, wrist fracture, vertebral fracture. In addition these values are age adjusted to allow for the decline in utility with age.

---

[58] A beta distribution of (1, 1) is characterized by a horizontal line where all probabilities are equally likely.

Values for normal health and fractures were estimated by direct elicitation from a sample of postmenopausal women (Cranney et al. 2001a)[59]. Preference scores for all health states were obtained through use of a visual analogue scale. In addition, both preference scores and utility values were obtained for the women's current health through the use of both a visual analogue scale and the standard gamble approach. Utility values for all health states were based on transforming preference scores based on estimating the power function relationship between preference scores and utility values for the study sample based on the data relating to current health (Torrance et al. 1995).

Health states describe the acute effects of fracture. Thus it is necessary to allow for the improvements in health status over time. Utility values for spine and wrist fractures are assumed to improve linearly so that the individual will be restored to normal health for their age group by the year end. Thus the total disutility from spine and wrist fractures will be half the difference between the utility value for normal health and fracture. For hip fracture, utility is assumed to improve linearly so that the individual will be restored to normal health for their age group by the end of the second year.

Utility decline by age is incorporated using data for women from the Canadian National Population Health Survey (Statistics Canada 1995b). The average age of respondents to the osteoporosis utility exercise was 55-59. Thus, the utilities for fracture states are converted to age specific utilities as by the following example:

---

[59] The study was conducted in Ottawa, Canada and was part of the MSc thesis of Ann Cranney for which I was supervisor.

$$Utility_{spine\ fracture\ aged\ 70-74} = \frac{Utility_{70-74}}{Utility_{55-59}} * Utility_{spine\ fracture}$$
$$= \frac{0.83}{0.86} * 0.77 = 0.74$$

Within the model the uncertainty around utility values is represented by a normal distribution with mean and standard errors derived from the utility elicitation exercise.

6.6 TREATMENT SPECIFIC PARAMETERS

6.6.1 Data Requirements

To facilitate the economic evaluation of treatment interventions, the model requires estimates of the following parameters for each of the therapies considered:

- Fracture specific relative risk reductions during therapy
- Benefit of therapy beyond therapy duration
- Continuation rate for therapy
- Monthly treatment costs

6.6.2 Effect of Therapy on Risk of Fractures

The effect of alternative bisphosphonates on the risks of fractures were derived from data from a meta analysis of randomized controlled clinical trials (Cranney et al. 2001b, 2002a, 2002b, 2002c). The meta analysis included randomized controlled trials of at least one year duration. Trials had to be of postmenopausal women with osteoporosis defined by prevalent fractures or low bone density (>2 standard deviations below the young adult mean).

The published meta analysis reports the relative risk of a woman having a vertebral or non vertebral fracture over the duration of the trial. For economic analysis, it is necessary to determine the decrease in the number of fractures not the number of women experiencing a fracture as well as distinguishing between wrist and hip fractures. In addition, as the model adjusts for actual compliance with therapy (rather than compliance in a trial setting) the relative risk reduction needs to relate to number of years of therapy and not number of years in the trial. Thus, each trial was reviewed and data was extracted based on the number of fractures and total years of therapy. Analysis focused on identifying the proportional risk reduction attributable to therapy for three specific fractures: hip, wrist and vertebral (Table 6.4).

6.6.3 Benefit of therapy beyond therapy duration

There is evidence that patients experience continued reductions in the risk of fracture after stopping therapy though follow up of patients has been for no more than two years post treatment curtailment (Tonino et al. 2000, Tosteson et al. 2001a). This analysis follows previous studies by assuming a linear reduction of benefit after the curtailment of therapy for all bisphosphonates though this benefit is restricted to a period of up to two years.

To incorporate uncertainty concerning the decline of benefit post treatment, the rate of decline of benefit was modeled as an exponential function where 0 represents a linear benefit, $-\varpi$ represents no decline over the two year period and $+\varpi$ represents no benefit post curtailment. This can be expressed by the formula:

Table 6.4: Meta Analysis for the Risk Reduction in Fractures

a. Alendronate

Fracture Site	Study	Alendronate Fractures	Alendronate Therapy Years	No therapy Fractures	No therapy Therapy Years	Relative Risk	(95% CI)
Hip	Black (1996)	11	2964	22	2915	0.492	(0.24, 1.01)
	Cummings (1998)	8	3440	18	3410	0.441	(0.19, 1.01)
	Liberman (1995)	1	1525	3	1015	0.222	(0.02, 2.13)
	Pols (1999)	2	950	3	958	0.672	(0.11, 4.01)
	Total	**22**	**8879**	**46**	**8298**	**0.465**	**(0.28, 0.77)**
Vertebral	Black	23	2964	50	2915	0.452	(0.28, 0.74)
	Chesnut (1995)	0	60	0	62	1.033	(0.02, 51.2)
	Cummings	22	3188	44	3188	0.500	(0.30, 0.83)
	Liberman	22	1525	40	1015	0.361	(0.22, 0.61)
	Total	**67**	**7737**	**134**	**7180**	**0.439**	**(0.33, 0.59)**
Wrist	Black	22	2964	41	2915	0.528	(0.32, 0.88)
	Cummings	34	3440	38	3410	0.887	(0.56, 1.41)
	Liberman	8	1525	16	1015	0.333	(0.14, 0.78)
	Pols	6	950	15	958	0.403	(0.16, 1.04)
	Total	**70**	**8879**	**110**	**8298**	**0.551**	**(0.35, 0.88)**

b. Etidronate

Fracture Site	Study	Etidronate		No therapy		Relative Risk (95% CI)	
		Fractures	Therapy Years	Fractures	Therapy Years		
Hip	Lyritis (1997)	1	179	2	163	0.455	(0.04, 4.97)
	Storm (1990)	1	76	2	74	0.487	(0.04, 5.26)
	Watts (1990)	3	407	0	397	6.828	(0.35, 131.7)
	Total	**5**	**662**	**4**	**634**	**0.945**	**(0.18, 4.91)**
Vertebral	Lyritis	10	179	19	163	0.479	(0.23, 1.00)
	Storm	18	76	43	74	0.408	(0.26, 0.64)
	Watts	12	407	25	397	0.468	(0.24, 0.92)
	Wimalawansa (1998)	3	56	5	56	0.600	(0.15, 2.39)
	Total	**43**	**718**	**92**	**690**	**0.443**	**(0.32, 0.61)**
Wrist	Lyritis	2	179	2	163	0.911	(0.13, 6.39)
	Storm	2	76	3	74	0.649	(0.11, 3.77)
	Watts	2	407	2	397	0.975	(0.14, 6.89)
	Total	**6**	**662**	**7**	**634**	**0.818**	**(0.28, 2.42)**

c. Risedronate

Fracture Site	Study	Risedronate		No therapy		Relative Risk	(95 % CI)
		Fractures	Therapy Years	Fractures	Therapy Years		
Hip	Harris (1999)	12	1953	15	1898	0.777	(0.36, 1.66)
	McLung (2001)	137	14253	95	7208	0.729	(0.56, 0.95)
	Mortensen (1998)	0	70	0	75	1.070	(0.02, 53.2)
	Reginster (2000)	9	954	11	920	0.789	(0.33, 1.90)
	Total	**158**	**17230**	**121**	**10101**	**0.739**	**(0.58, 0.94)**
Vertebral	Clemmesen (1997)	13	73	20	75	0.668	(0.36, 1.24)
	Harris	61	1619	93	1712	0.694	(0.51, 0.95)
	Mortensen	1	70	0	75	3.211	(0.13, 77.6)
	Reginster	53	879	89	921	0.624	(0.45, 0.87)
	Total	**128**	**2641**	**202**	**2783**	**0.665**	**(0.54, 0.82)**
Wrist	Harris	14	1953	22	1898	0.618	(0.32, 1.20)
	Mortensen	0	70	0	75	1.070	(0.02, 53.2)
	Reginster	15	954	21	920	0.689	(0.36, 1.34)
	Total	**29**	**2977**	**43**	**2893**	**0.658**	**(0.41, 1.05)**

$$RR_{fracture} = \left(\frac{3 - T_{Stop}}{3} \right)^{e^r} \bullet RR_{treatment}$$

T_{stop} = years since curtailment of therapy

$RR_{fracture}$ = relative risk of fracture

$RR_{treatment}$ = relative risk of fracture on therapy

r = the rate of decline of benefit after curtailment, =0 for a linear reduction in benefit

Uncertainty around the rate of decline of benefit was characterized by a normal distribution (μ=0, σ =1). The extremes associated with the distribution correspond to the two scenarios suggested by the WHO Collaborating Centre (1999); that all benefit is lost immediately or that the relative risk of fracture remains as estimated in the trial. Thus, as with the attributable fracture of long term care stay (Section 6.4.5) the model adopts an assumption that the true proportion is uncertain and lies between these extremes.

6.6.4 Continuation with Therapy

Retention rates for bisphosphonates were estimated using data from the Ontario Drug Benefit Scheme. Data related to claimants new to each drug over a period from July 1999 to June 2001. One year retention rates were obtained for each drug based on the proportion of claimants new to the drug that continued on therapy after one year. Rates for alendronate, risedronate and etidronate were 65%, 62% and 57% respectively. Thus, continuation with therapy is assumed to decline yearly based on the observed retention rates. Duration is assumed to be no more than seven years in line with the maximum follow up of patients in clinical trials (Tonino et al.2000).

However, it should be noted that based on the above the proportion of patients who are likely to be continuing with therapy at seven years will be between 2 and 5%.

Despite the large sample sizes (n=115 426) used to estimate retention rates, uncertainty is assumed to be high as observed data relate to the first twelve months on therapy. Thus, to more accurately reflect the uncertainty concerning long term continuation with therapy, continuation rates were characterized by beta distributions with a presumed sample size of 100.

6.6.5 Cost of therapies

The costs of alendronate, etidronate and risedronate were obtained from the Ontario Drug Benefit Plan (Ontario Ministry of Health and Long Term Care 2002). Pharmacy mark up and dispensing fees were obtained from local pharmacies. The yearly cost of alendronate was estimated to be $730.59 based on a dosage of 10 mg once daily incorporating a 10% mark up and a $6.41 dispensing fee. The yearly cost of etidronate based on a dosage of 200 mg twice daily for 14 days in a three month cycle was estimated to be $189.72. The yearly cost of risedronate based on a dosage of 5 mg once daily was estimated to be $692.49. Costs of drug therapies are assumed fixed.

6.7 QUALITY OF THE ECONOMIC MODEL

The previous chapter included a review of guidance on the conduct of economic evaluation in osteoporosis. Four guidance documents were identified (Coyle and Tosteson 2003, Tosteson et al. 2001a, Dere et al. 1998, WHO Collaborating Centre 1999). Based on the recommendations within these documents, fifteen

methodological issues of importance with respect to economic models for osteoporosis were identified. The economic model of osteoporosis outlined above corresponds to the recommendations made for each of these issues as outlined in Table 6.5.

6.8 ANALYTICAL FRAMEWORK

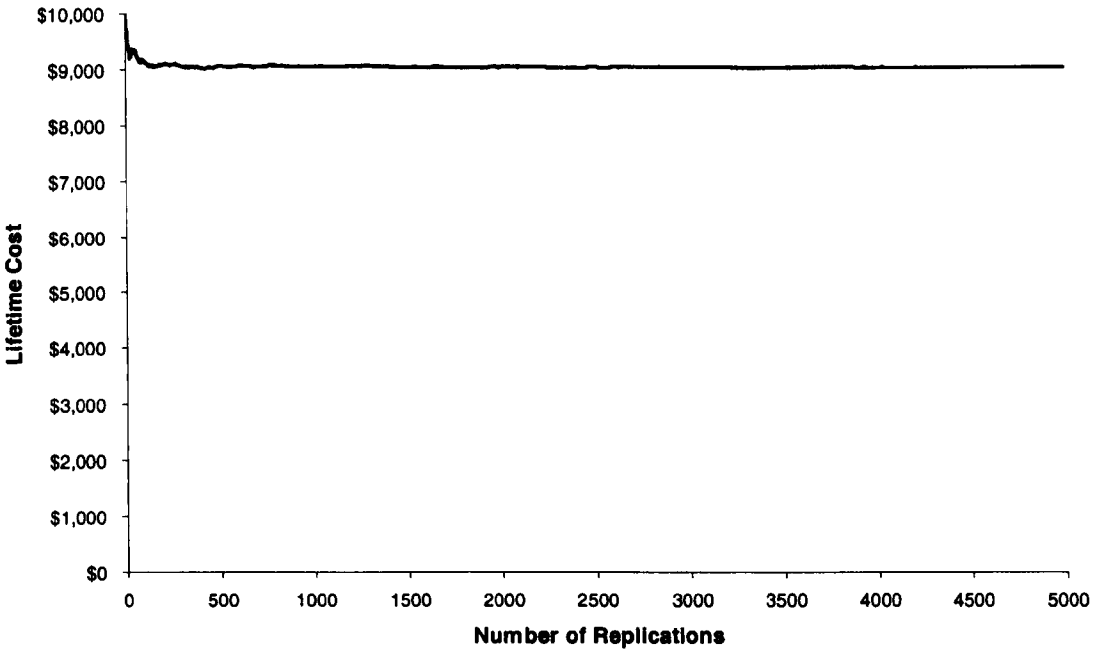
The model is designed to allow the conduct of a cost utility analysis. Analysis presented in Chapter 7 and 8 is conducted from the perspective of a provincial ministry of health. The model adopts a lifetime horizon with costs and benefits discounted at 5% per annum (CCOHTA 1997). All costs are presented in terms of 2001 Canadian dollars. The model estimates expected values for costs and QALYs for 62 patient cohorts characterized by the patient's age (65-95) and fracture history (no previous osteoporotic fracture and previous osteoporotic fracture).

Analysis is conducted through Monte-Carlo simulation whereby estimates of outcomes are obtained by repeated sampling from the probability distribution of input parameters. Larger numbers of replications provide more stable estimates of expected values. The decision model is complex given that costs and effects are modeled for 62 patient cohorts. Thus, it takes approximately 1 ½ hours for each 1000 replications. Based on the uncertainty over the expected value of various outcomes, simulations with 3000 replications are argued to be appropriate (Figure 6.4). The adoption of the Monte Carlo simulation facilitates an optimal treatment choice as it provides a more accurate expression of the expected value of outcomes

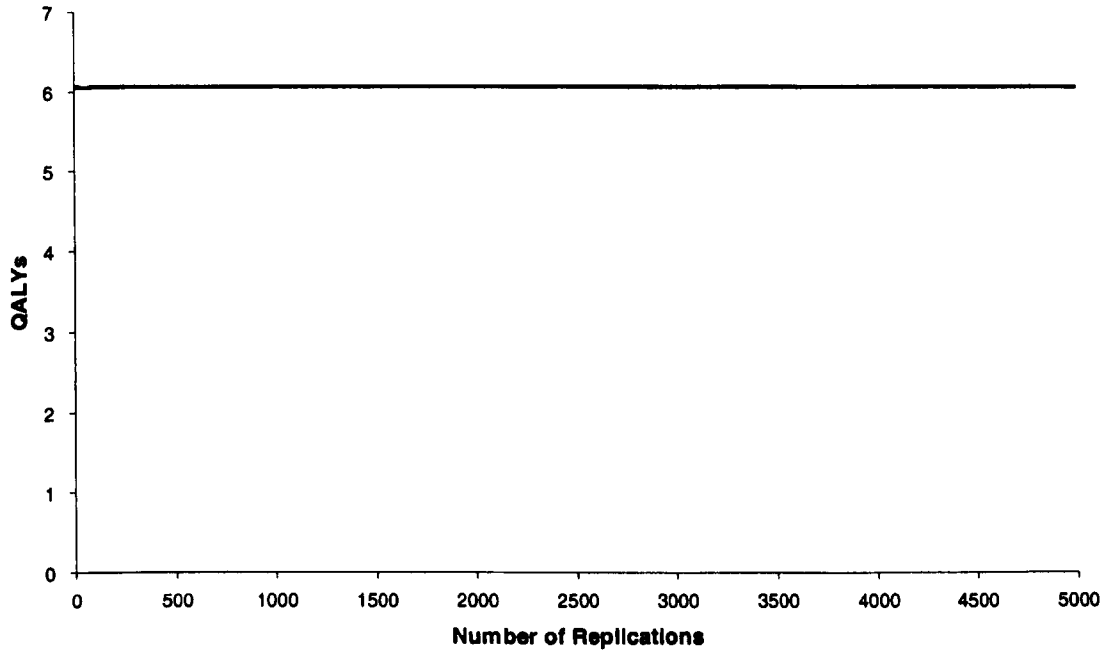
Table 6.5: How Methodological Issues are Addressed within the Model

Methodological Issue	How addressed
Time horizon	Model incorporates all long term effects with a lifetime horizon
Study perspective	Perspective is that of the health care system
Modelling fractures	Model uses age-specific fracture incidence with treatment effects modelled using relative risk reductions from a meta analysis of RCTs
Data sources	All data are specific to Canada except the relative risk reductions and the risk of mortality post vertebral fracture
Mortality following fractures	Model includes attributable mortality following hip and vertebral fracture
Long term care admission post fractures	Analysis assumes that attributable long term care is uncertain but is bounded by all future long term care costs being attributable to fractures and that only first year costs being attributable
Head to head comparisons	Is achieved through synthesis of similarly designed trials.
Extraskeletal effects	All effects which impact cost effectiveness are included. No significant adverse effects associated with bisphosphonates
Benefit beyond therapy	Analysis assumes that benefit beyond therapy is uncertain but bounded by there being no benefit beyond therapy and that benefit is maintained for a period of up to two years post discontinuation.
Model validation	Model is fully calibrated replicating all population data
Compliance with therapy	Discontinuation rates based on Ontario Drug Benefit formulary data
Utility values	Utility values vary by fracture state and age
Accommodating variability	Analysis is stratified by age and fracture history.
Accommodating uncertainty	Probabilistic analysis is conducted to determine optimal therapies and to determine optimal future research.
Study funding	Study is not funded by industry and is part of PHD thesis: previous funding for the development of the model has been given by Merck Canada and Novartis Canada

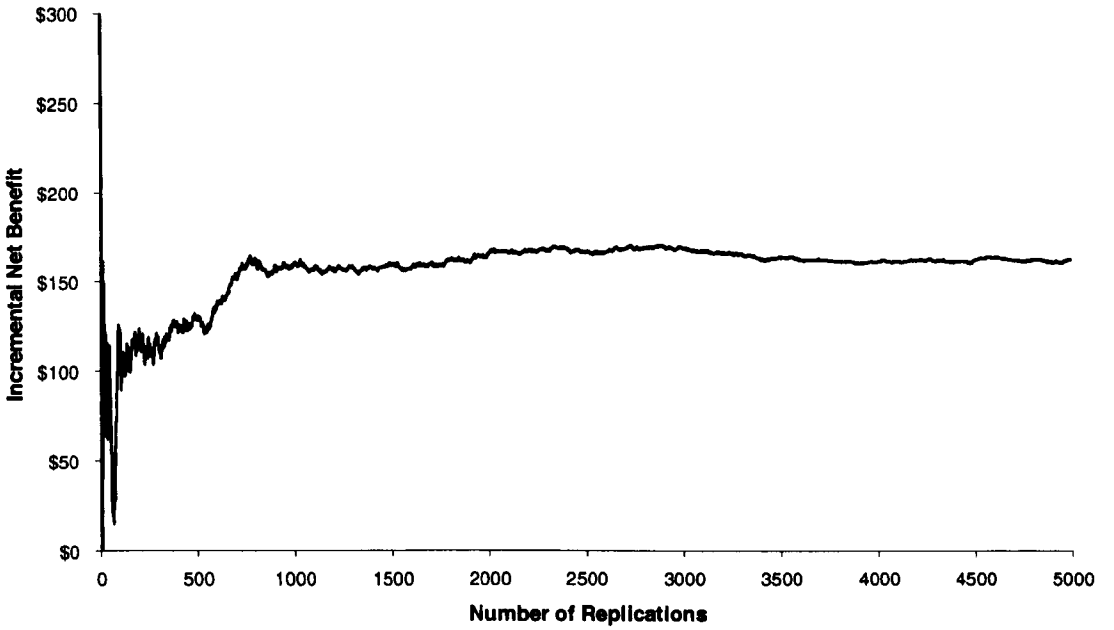
Figure 6.4: Expected Value of Outcome Parameters Based on a 75 Year-old Woman with Previous Fracture as a Function of Number of Replications



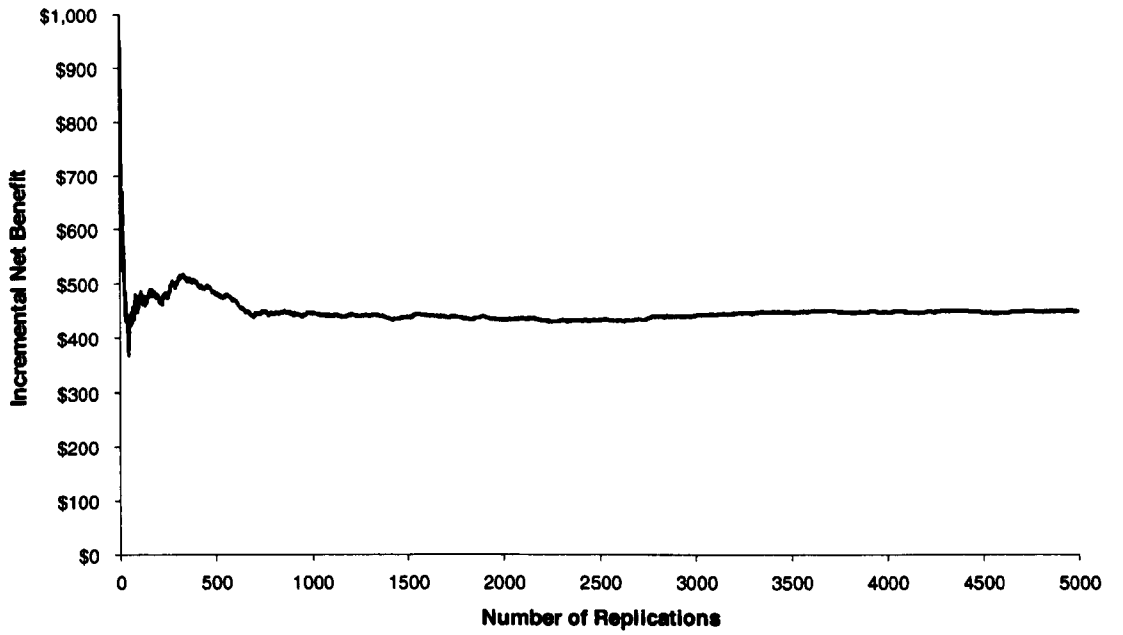
a. Lifetime Costs with Alendronate



b. QALYs with Alendronate



c. Net Benefit of Alendronate compared to No therapy



d. Global EVPI

of interest than deterministic analysis given the uncertainty propagated by parameter estimates (Thompson and Graham 1996).

As per previous chapters, cost effectiveness can be expressed in terms of both the incremental cost per QALY (ICER) gained and the incremental net monetary benefit (INB) (Stinnett and Mullahy 1998). The focus of the analysis will be to determine an optimal prescribing policy with respect to which treatments would be appropriate for which patients (Chapter 7)[60]. Value of information analysis will determine optimal research designs aimed to reduce decision uncertainty (Chapter 8).

6.9 DISCUSSION

In this chapter, a decision model which facilitates an economic analysis of treatment options for the prevention of osteoporotic fractures is detailed. The model adheres to the 15 recommendations for the conduct of such analyses detailed in Chapter 5. The model incorporates Canadian data on the risks, costs and disutilities associated with osteoporotic fractures. The model allows the conduct of both probabilistic analysis through the specification of probability density functions for relevant input parameters and stratified analysis based on a woman's age and previous fracture history.

In Chapter 7, the results of the stratified analysis are presented. The results are based on expected values obtained form the probabilistic analysis. The chapter contains recommendations relating to the appropriate limited use criteria for the reimbursement of bisphosphonates within Ontario. In Chapter 8, a value of

[60] The same data source used for estimation of continuation rates on therapy is used to determine the prior breakdown of current claimants receiving bisphosphonates by age and drug (see Figure 7.1).

information analysis is presented whereby, the uncertainty concerning the individual parameters within the decision model is considered and an optimal plan for further research relating to such uncertainty is identified.

Chapter 7.

Stratified Cost Effectiveness Analysis of Treatment for Osteoporosis in Canada

---

7.1 INTRODUCTION

In Chapter 3 a framework for handling variability with respect to the economic evaluation of health care interventions was identified. Variability relates to the lack of homogeneity in outcomes amongst the target patient population. The framework provides an intuitive approach to handling such heterogeneity through stratified cost effectiveness analysis. Stratified analysis can assist in determining appropriate limited use criteria for interventions. Through stratified analysis patients are placed into groups which are more homogenous with respect to cost effectiveness. The results of a stratified analysis facilitate the development of limited use criteria (LUC) which allow technologies to be restricted to specific patient populations.

In Chapter 5, recommendations for the conduct of economic analysis in osteoporosis were identified. In Chapter 6, a decision model for the conduct of economic evaluations in osteoporosis for Canada was described in detail. This chapter presents the results of a stratified cost effectiveness analysis for the treatment of osteoporosis by bisphosphonates using this model. In Section 7.2, the background to the conduct of the analysis and its subsequent use in decision making is provided. In Section 7.3, the methods of the analysis are detailed. In Section 7.4, the results are presented.

7.2 BACKGROUND

In Canada, individual provinces have the responsibility for determining the level that prescription drug costs are subsidized for different individuals and which prescription drugs are subsidized. In Ontario, provincial drug coverage is controlled by the Ministry of the Health and Long Term Care through the Ontario Drug Benefit Formulary (ODB) (Ontario Ministry of Health and Long Term Care 2002). Coverage is provided both for elderly (aged over 65) and for low income families. Decisions on which drugs are covered within the provincial formulary are taken through consideration of recommendations made by the Drug Quality and Therapeutics Committee (DQTC) which comprises both clinicians and pharmacists (Laupacis 2002). The DQTC considers both clinical and economic evidence provided by the pharmaceutical company requesting coverage of their product (Ontario Ministry of Health and Long Term Care 2000). Most drugs are given limited access whereby drug costs are covered based on criteria more restrictive than that for which the drug is licensed[61,62] (Ontario Ministry of Health and Long Term Care 2004).

In January 2003, of the three bisphosphonates licensed for the treatment of osteoporosis only etidronate was unrestricted with respect to coverage through the Ontario Drug Benefit Formualry (Ontario Ministry of Health and Long Term Care 2002). Access to alendronate and risedronate was restricted to individuals who were categorized as failures with respect to therapy with etidronate. Failure was based on either an increase in bone mineral density (BMD) after 2 years with etidronate, the

---

[61] Between 1999 and 2001, 20 of 37 drugs approved for coverage through the Ontario Drug Benefit Formulary were subject to criteria more restrictive than the products licensing.

[62] This process is more formal than that of the National Institute of Clinical Excellence in the UK.

incidence of a clinical fracture or failure to tolerate etidronate. There was general recognition for the need to revise the limited use criteria for bisphosphonates. A previous version of the following analysis was instrumental in the revisions adopted in the spring of 2003 (Ontario Ministry of Health and Long Term Care 2003). The following analysis incorporates further enhancements to the model though the conclusions and recommendations are similar to the original analysis.

7.3 ANALYSIS PLAN

7.3.1 Net Benefit Gain from Stratification

Analysis uses the Canadian Economic Model for osteoporosis outlined in Chapter 5. The model is used to derive expectations for the lifetime cost and QALYs associated with each of four treatment options; no therapy, etidronate, alendronate and risedronate. Expected values are obtained through Monte Carlo simulation with 3000 replications.

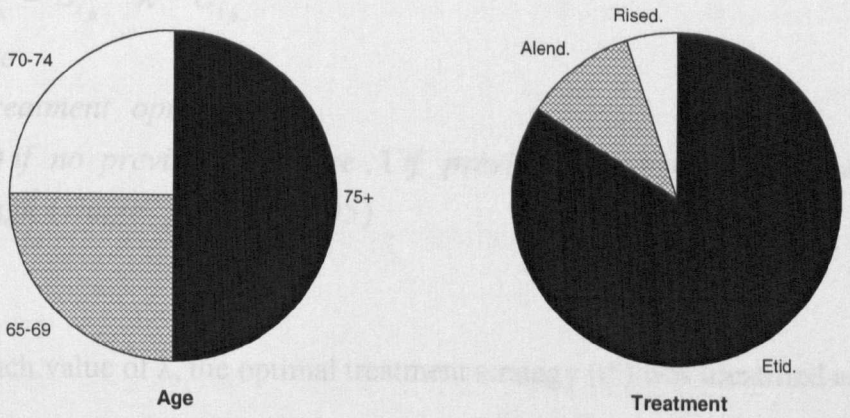
Analysis is conducted for 62 strata for osteoporotic women determined by both age (31 age groups from 65 to 95[63]) and previous history of osteoporotic fracture (no previous fracture and previous fracture)[64]. The number of women in each strata who are potential patients were derived through examination of the Ontario Drug Benefit Formulary data for 2001 (Figure 7.1). The majority of women were treated with etidronate (74%) with an almost equal split between women aged over 75 (48%) and under 75 (52%).

---

[63] Analysis is restricted to women aged over 65 as this is the cut off age for all to receive benefits under the Ontario Drug Benefit Formulary. Only a relatively small number of women under 65 who receive social assistance may require treatment for osteoporosis.

[64] Other bases of stratification could have been considered. For instance, the risk of hip fracture is greater for women who consume a high degree of alcohol. It was decided that this was an unlikely basis from which to discriminate amongst potential patients. The two bases chosen were practical, acceptable and had the greatest degree of literature demonstrating a link with fractures.

Figure 7.1: Distribution of New Claimants (n=115 426) from the Ontario Drug Benefit Plan for Bisphosphonates in 2001



Source: Ontario Drug Benefit Formulary data for 2001

The net benefit gain from stratification is determined based on the methods outlined in Chapter 2. First, the net benefit for each stratum was estimated for a range of values of λ.

$$NB_{t_{jk}} = E_{t_{jk}} * \lambda - C_{t_{jk}}$$

where

t = *treatment option*

$j = 0$ *if no previous fracture*, 1 *if previous history of fracture*

k = *age (range from* 65 *to* 95)

For each value of λ, the optimal treatment strategy (t*) was identified assuming there was no stratification:

$$TNB_t = \sum_{j=0,1\ k=65,66..95} NB_{t_{jk}} * n_{jk}$$

$$TNB_{t^*} = \max_t (TNB_t)$$

where

n = *number of women in each strata*

With stratification, the optimal treatment strategy for each strata (t_{jk}*) were identified and the total net benefit (TNB_{sjk}) derived.

$$NB_{t_{jk}*} = \max_t (NB_{jkt})$$

$$TNB_{sjk} = \sum_{j=0,1\ k=65,66..95} NB_{t_{jk}*} * n_{jk}$$

From the above, optimal limited use criteria for each treatment option were derived for alternate values of λ, by determining the optimal treatment choice (that which maximizes net benefit) for each of the 62 strata.

The total net benefit gain from stratification ($\Delta_S TNB$) for each value of λ was thus:

$$\Delta_S TNB = TNB_{sjk} - TNB_{t*}$$

7.3.2 Net Benefit Loss from Restricted Stratification

Stratification could be restricted to either only a woman's age or their fracture history. This could occur either for logistical reasons, in that it may be difficult to enforce stratification using a particular basis, or for concerns for equity. Restriction of stratification to only age or fracture history can lead to a net benefit loss (Δ_ETNB).

For example, the net benefit loss from restricting stratification to fracture history only could be estimated as follows:

$$\Delta_E TNB = TNB_{sjk} - TNB_{sj}$$

where

$$TNB_{sj} = \sum_{j=0,1} \max_t \left(NB_{jt}\right)^* .n_j$$

The net benefit loss was estimated for alternate values of λ for both restricting stratification to age only and to fracture history only.

7.3.3 Impact of Leakage

Analysis of leakage focused solely on indiscriminant leakage[65].

If, under certain values of λ, a bisphosphonate was the optimal treatment choice without stratification, leakage can not lead to a change in the optimal limited use criteria - rather leakage would lead to a reduction in the net benefit gain from

[65] No a priori hypotheses relating to discriminant leakage were considered.

stratification. Thus, for those values of λ, analysis focused on the net benefit loss from leakage (Δ_LTNB) given different levels of leakage (l).

$$\Delta_L TNB = \sum_{j=0,1\ k=65,66..95} \left(NB_{t_{jk}*} - NB_{t*}\right) l.n_{jk}$$

However, for values of λ where no therapy was the optimal treatment choice without stratification, it is possible that leakage may be such that the optimal limited use criteria would be to deny treatment to all strata. That is, the net benefit loss from leakage will be greater than the net benefit gain from allowing limited access to therapy. Thus for those values of λ, a threshold level of leakage is identified which corresponds to the level of leakage required to make such a reverse in decision optimal: i.e. where the net benefit loss from leakage is equivalent to the net original net benefit gain from stratification.

$$l^* = \frac{\sum_{j=0,1\ k=65,66..95} \left(NB_{t_{jk}*} - NB_{t*}\right) n_{jk}}{TNB_{sjk}}$$

7.3.4 Impact of Budget Constraints

The analysis above assumes that if the optimal limited use criteria required an increase in either the overall health care budget or the budget of the Ontario Drug Benefit Formulary then this would be acceptable given the increase in net benefit. However, if the budget for either health care in total or for drugs alone is fixed then an appropriate LUC should maximize the net benefit gain from stratification subject to the relevant budget constraint.

The current lifetime osteoporosis related health care and drug expenditures in Ontario for patients newly prescribed bisphosphonates was estimated by taking

prescribing rates from the Ontario Drug Benefit Formulary and forecasting expenditures with the decision model. These costs were assumed to be the relevant budget constraints for health care (B_h) and drugs (B_d).

Thus, analysis identified the maximum net benefit gain from stratification based on these constraints ($TNB_s|B$) and the net benefit loss arising from enforcement of the budget constraint ($\Delta_B TNB$).

$$\Delta_{B_h} TNB = TNB_{sjk} - TNB_{sjk}|B_h$$
$$\Delta_{B_d} TNB = TNB_{sjk} - TNB_{sjk}|Bd$$

7.3.5 Comparison of Optimal Limited Use Criteria with Revised Limited Use Criteria

Analysis described above focused on the net benefit gain from using stratification in determining an optimal prescribing policy compared to no stratification. However, the Ontario Drug Benefit Formulary LUC for bisphosphonates at the beginning of 2003 was not based upon either of these frameworks. Thus, further analysis examined the net benefit gain ($\Delta_s TNB|2003$) when comparing the current prescribing practice to the optimal limited use criteria determined by the analysis in Section 7.3.1.

$$\Delta_s TNB|2003 = TNB_{sjk} - TNB_{2003}$$
$$TNB2003 = \sum_{j=0,1\,k=65,66..95\,t=1,2,3} n_{jkt} * NB_{jkt}$$

where

t = *treatment* (1 = *alendronate*, 2 = *etidronate*, 3 = *risedronate*)

n_{jkt} = *number of women newly prescribed drug under the ODB formulary*

7.4 RESULTS

7.4.1 Net Benefit Gain from Stratification

Table 7.1 details the expected values for the lifetime costs and QALYs associated with all four treatment strategies for each of the 62 strata.

Table 7.2 provides the expected values for lifetime costs and QALYs for each treatment strategy if the strategy was given to all patients: i.e. the expected values from Table 7.1 are weighted by proportion of patients in each stratum. Etidronate is dominated by both alendronate and no therapy whilst risedronate is dominated by alendronate. If the value of a QALY is less than $37 000 then without stratification no therapy is the optimal treatment choice, otherwise alendronate is optimal (Table 7.2).

Figure 7.2 presents the incremental costs and QALYs from each of the three bisphosphonates compared to no therapy for all strata. For alendronate and risedronate, the incremental costs of treatment are smaller for older patients and for patients with previous fracture whilst the incremental QALYs gained tend to be larger for both groups[66]. For etidronate, the higher the risk of fracture the higher the incremental costs and the lower the incremental QALYs.

For all strata, alendronate and no therapy dominate etidronate and alendronate dominates risedronate. Thus, for all values of λ the optimal limited use criteria will include only alendronate or no therapy. At a value of λ of $50 000, the optimal stratification would be to restrict alendronate to women with previous fracture aged

[66] Note that for the oldest patient groups, incremental QALYs begin to decrease by age as the capacity to gain from avoided hip fractures falls.

Table 7.1: Expected Lifetime Costs and QALYs by Age, Fracture History and Treatment Strategy

a. Previous fracture history

Age	No Therapy		Alendronate		Etidronate		Risedronate	
	Costs	QALYs	Costs	QALYs	Costs	QALYs	Costs	QALYs
65	$6,193	9.490	$7,436	9.497	$6,650	9.488	$7,387	9.493
66	$6,390	9.173	$7,609	9.180	$6,857	9.170	$7,572	9.176
67	$6,592	8.846	$7,783	8.853	$7,070	8.843	$7,759	8.849
68	$6,800	8.510	$7,958	8.518	$7,290	8.506	$7,950	8.514
69	$7,013	8.164	$8,133	8.173	$7,517	8.160	$8,144	8.168
70	$7,232	7.808	$8,306	7.818	$7,752	7.803	$8,341	7.812
71	$7,456	7.471	$8,475	7.482	$7,999	7.466	$8,537	7.476
72	$7,688	7.126	$8,636	7.138	$8,262	7.121	$8,735	7.132
73	$7,931	6.773	$8,793	6.786	$8,544	6.767	$8,935	6.779
74	$8,187	6.412	$8,934	6.426	$8,858	6.405	$9,133	6.418
75	$8,458	6.041	$9,062	6.056	$9,205	6.033	$9,333	6.048
76	$8,516	5.704	$9,060	5.721	$9,286	5.696	$9,358	5.712
77	$8,566	5.360	$9,047	5.378	$9,359	5.351	$9,374	5.369
78	$8,609	5.010	$9,023	5.029	$9,426	5.000	$9,381	5.019
79	$8,647	4.651	$8,988	4.672	$9,488	4.641	$9,379	4.661
80	$8,680	4.283	$8,940	4.304	$9,545	4.272	$9,368	4.293
81	$8,702	3.999	$8,881	4.022	$9,597	3.987	$9,346	4.009
82	$8,726	3.710	$8,811	3.734	$9,655	3.698	$9,318	3.722
83	$8,756	3.417	$8,734	3.442	$9,724	3.404	$9,290	3.429
84	$8,797	3.117	$8,643	3.143	$9,818	3.103	$9,260	3.129
85	$8,853	2.806	$8,539	2.834	$9,943	2.792	$9,230	2.819
86	$8,706	2.642	$8,320	2.672	$9,813	2.626	$9,038	2.656
87	$8,543	2.483	$8,087	2.514	$9,665	2.466	$8,831	2.497
88	$8,365	2.329	$7,840	2.362	$9,501	2.311	$8,609	2.344
89	$8,171	2.180	$7,577	2.215	$9,317	2.162	$8,371	2.196
90	$7,956	2.037	$7,298	2.074	$9,109	2.018	$8,112	2.054
91	$7,715	1.900	$6,996	1.938	$8,873	1.880	$7,829	1.918
92	$7,442	1.768	$6,669	1.807	$8,601	1.747	$7,516	1.786
93	$7,129	1.642	$6,313	1.682	$8,283	1.621	$7,166	1.661
94	$6,759	1.522	$5,917	1.562	$7,902	1.501	$6,767	1.541
95	$6,306	1.408	$5,473	1.446	$7,423	1.387	$6,301	1.426

b. No previous fracture

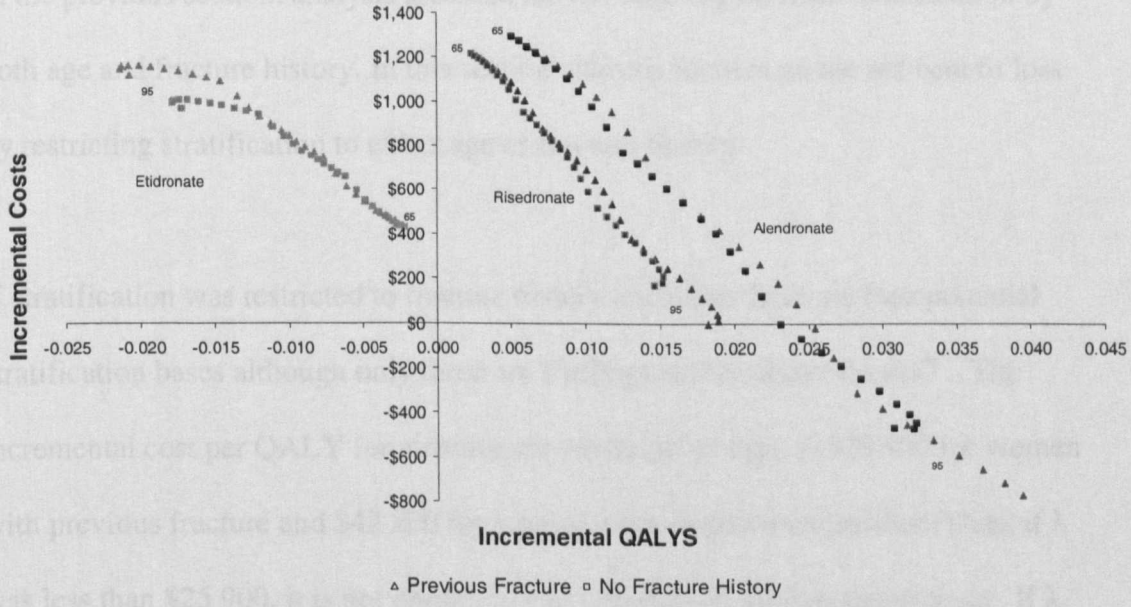
Age	No Therapy		Alendronate		Etidronate		Risedronate	
	Costs	QALYs	Costs	QALYs	Costs	QALYs	Costs	QALYs
65	$5,010	9.517	$6,299	9.522	$5,435	9.515	$6,224	9.520
66	$5,167	9.200	$6,435	9.206	$5,600	9.198	$6,369	9.203
67	$5,327	8.874	$6,572	8.881	$5,770	8.872	$6,517	8.877
68	$5,492	8.539	$6,709	8.546	$5,944	8.536	$6,667	8.542
69	$5,661	8.194	$6,845	8.202	$6,124	8.191	$6,819	8.198
70	$5,835	7.838	$6,979	7.847	$6,310	7.835	$6,973	7.842
71	$6,009	7.503	$7,108	7.512	$6,503	7.499	$7,124	7.507
72	$6,189	7.159	$7,229	7.169	$6,707	7.155	$7,274	7.163
73	$6,376	6.806	$7,347	6.817	$6,926	6.802	$7,426	6.811
74	$6,572	6.446	$7,451	6.457	$7,168	6.440	$7,576	6.451
75	$6,779	6.076	$7,545	6.088	$7,434	6.070	$7,725	6.082
76	$6,824	5.739	$7,539	5.753	$7,497	5.733	$7,742	5.746
77	$6,863	5.396	$7,524	5.411	$7,555	5.389	$7,750	5.403
78	$6,896	5.046	$7,499	5.062	$7,608	5.038	$7,752	5.053
79	$6,926	4.688	$7,465	4.704	$7,657	4.680	$7,746	4.6
80	$6,951	4.320	$7,421	4.337	$7,701	4.311	$7,732	4.328
81	$6,963	4.036	$7,363	4.054	$7,738	4.027	$7,704	4.045
82	$6,975	3.747	$7,295	3.767	$7,779	3.738	$7,672	3.757
83	$6,991	3.454	$7,222	3.475	$7,828	3.444	$7,638	3.464
84	$7,014	3.154	$7,136	3.176	$7,895	3.143	$7,601	3.164
85	$7,047	2.843	$7,040	2.867	$7,983	2.831	$7,562	2.854
86	$6,924	2.679	$6,855	2.704	$7,876	2.666	$7,400	2.691
87	$6,789	2.520	$6,658	2.545	$7,755	2.506	$7,225	2.532
88	$6,642	2.365	$6,450	2.392	$7,620	2.351	$7,038	2.378
89	$6,481	2.217	$6,230	2.245	$7,470	2.201	$6,837	2.230
90	$6,304	2.073	$5,995	2.103	$7,300	2.057	$6,620	2.087
91	$6,105	1.935	$5,742	1.966	$7,107	1.918	$6,382	1.950
92	$5,881	1.802	$5,469	1.834	$6,884	1.785	$6,121	1.817
93	$5,623	1.675	$5,173	1.707	$6,623	1.657	$5,829	1.690
94	$5,319	1.554	$4,845	1.586	$6,309	1.536	$5,499	1.569
95	$4,951	1.437	$4,480	1.467	$5,916	1.420	$5,115	1.451

Table 7.2: Average Lifetime Costs and QALYs by Treatment Strategy without Stratification

	No Therapy	Alendronate	Etidronate	Risedronate
Lifetime Costs	$7,310	$7,908	$8,018	$8,163
QALYs	5.991	6.007	5.983	5.998
ICER versus no Therapy		$37 000	Dominated	$112 600

Note: Data are obtained by weighting the expected value of costs and QALYs for each cohort in Table 7.1 by the proportion of claimants in each cohort.

Figure 7.2 Incremental Costs and QALYs for Bisphosphonates Compared to No Therapy for Each Strata



75 or greater and for osteoporotic women without previous fracture aged 77 or greater (Figure 7.3). For all other women, no therapy will be optimal. The greater the value of λ, the younger the cut off age point for access to alendronate (Figure 7.3). Hence, the greater the value of λ, the greater the proportion of women for whom alendronate should be made available (Figure 7.4).

Figure 7.5 shows that at a value of λ of $50 000, the net benefit from stratification is $24.9 million when compared to the optimal policy without stratification which would be to make alendronate available to all women. The net benefit from stratification varies by λ with a peak at $37 000 (Figure 7.5). For low values of λ, no therapy will be optimal for most women thus there is little value from stratification. Similarly for high values, alendronate will be optimal for most women.

7.4.2 Net Benefit Loss from Restricted Stratification

In the previous section, analysis assessed the net benefit gain from stratification by both age and fracture history. In this section analysis focuses on the net benefit loss by restricting stratification to either age or fracture history.

If stratification was restricted to fracture history only then there are four potential stratification bases although only three are likely given the above results[67]. The incremental cost per QALY for alendronate versus no therapy is $25 900 for women with previous fracture and $42 200 for women without previous fracture. Thus, if λ was less than $25 900, it is not optimal for any women to receive alendronate. If λ

[67] The four potential bases are alendronate available to no women, alendronate restricted to those with previous fracture only, alendronate restricted to women with no previous fracture and alendronate available to all. Obviously given the increased benefit of alendronate for women with previous fracture, it is unlikely that a policy of restricting alendronate to only women with no fracture history would be optimal.

Figure 7.3: Age Cut Offs for Restricted Access to Alendronate by Value of a QALY

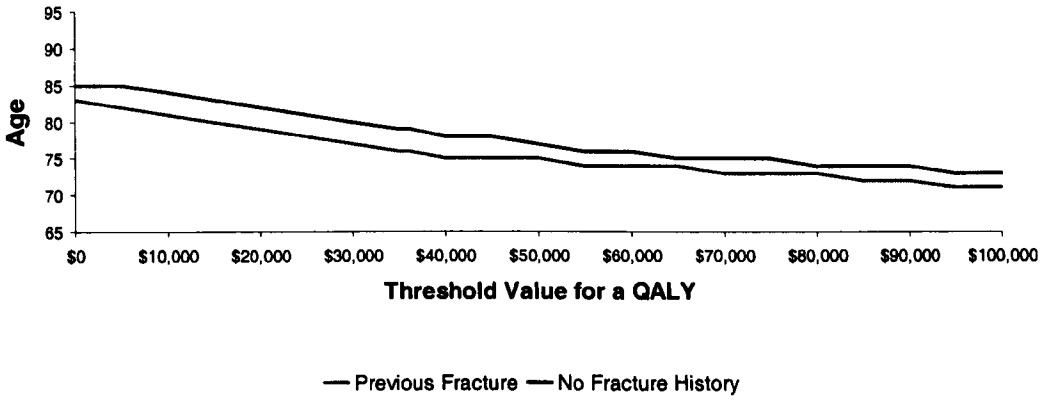


Figure 7.4: Percentage of Women for whom Alendronate is Optimal by Value of a QALY

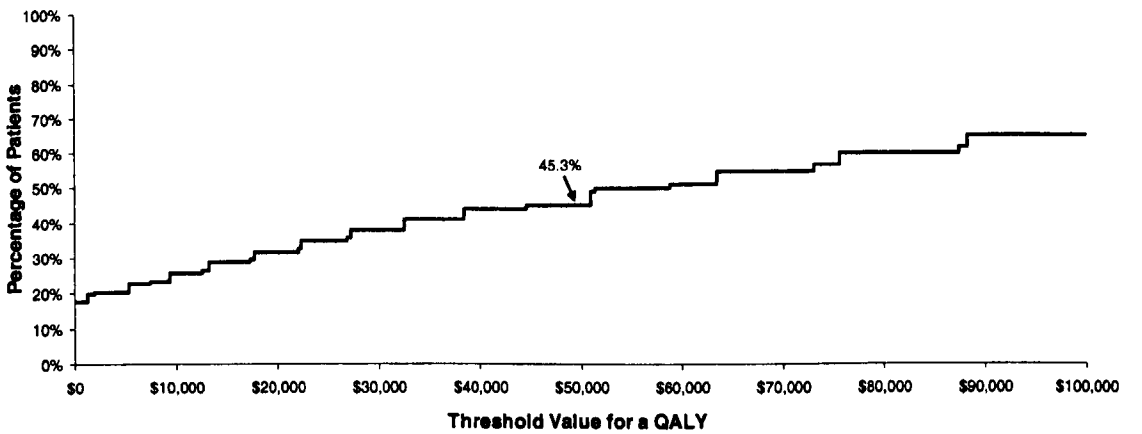
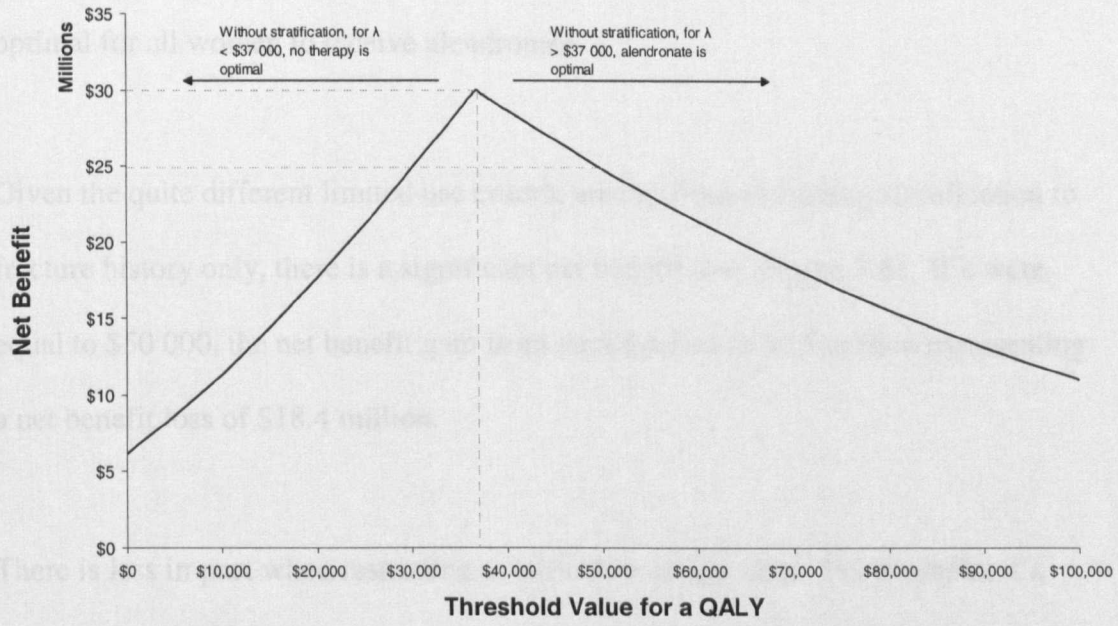


Figure 7.5: Net Benefit from Stratification by Value of a QALY



was valued between $25 900 and $72 900, it is optimal for women with previous fracture to receive alendronate (Table 7.3). For values of λ greater than $72 900 it is optimal for all women to receive alendronate.

Given the quite different limited use criteria arising from restricting stratification to fracture history only, there is a significant net benefit loss (Figure 7.6). If λ were equal to $50 000, the net benefit gain from stratification is $6.5 million representing a net benefit loss of $18.4 million.

There is less impact when restricting stratification to age only. For example, if λ were equal to $50 000, the optimal limited use criteria when stratifying by age only is to restrict alendronate to women aged 75 or greater. Thus, the impact of not stratifying by fracture history is that alendronate would now be available for women aged 75 or 76. The net benefit loss from the revised stratification basis is $130 000.

7.4.3 Impact of Leakage

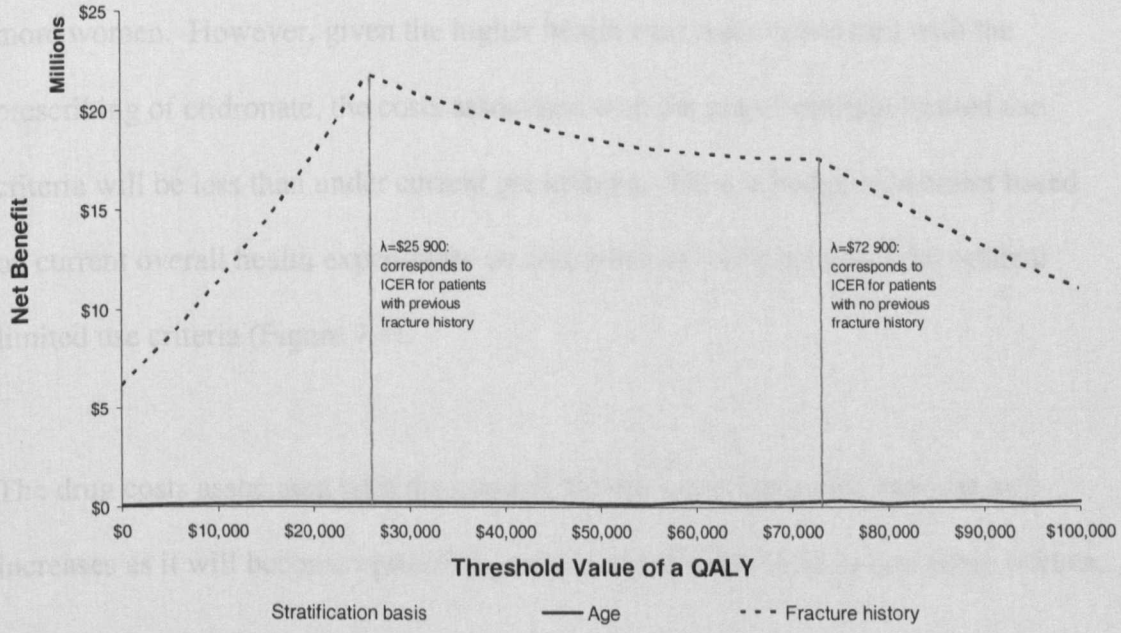
For values of λ more than $37 000, the impact of leakage would be to reduce the net benefit gain from stratification. For example, if λ were equal to $50 000 and the level of leakage was 10%, the net benefit loss from leakage is $2.49 million ($24.9 million *10%).

For values of λ less than $37 000, there is the potential that leakage may lead to a revised stratification bases whereby alendronate would not be made available to any women. This occurs when the bet benefit gain from making alendronate within a small number of strata is less than the loss from leakage to other strata.

Table 7.3: Average Lifetime Costs and QALYs by Treatment Strategy Stratified by Fracture History

	No Therapy	Alendronate	Etidronate	Risedronate
No previous fracture				
Lifetime Costs	$6 103	$6 965	$6 695	$7 089
QALYs	6.701	6.713	6.695	6.706
ICER versus no Therapy		$72 900	Dominated	$177 600
Previous fracture				
Lifetime Costs	$7,877	$8,351	$8,640	$8,667
QALYs	5.657	5.675	5.648	5.666
ICER versus no Therapy		$25 900	Dominated	$92 700

Figure 7.6: Net Benefit Loss from Restricted Stratification by Value of a QALY



For example, if λ was equal to $20000, the net benefit gain from stratification is $17.6 million. The incremental net benefit for no therapy versus alendronate in those strata where no therapy is optimal is $38 million. Thus, if the level of leakage is greater than 46% then a policy of not allowing alendronate to any women becomes optimal[68].

This break even point for the level of leakage is greater the value of λ (Figure 7.7).

7.4.4 Impact of Budget Constraints

Based on the incidence rates for prescribing bisphosphonates and the estimated costs from the decision model, the lifetime osteoporosis related health care and drug costs associated with individuals newly prescribed bisphosphonates would be $600.4 million and $41.8 million respectively.

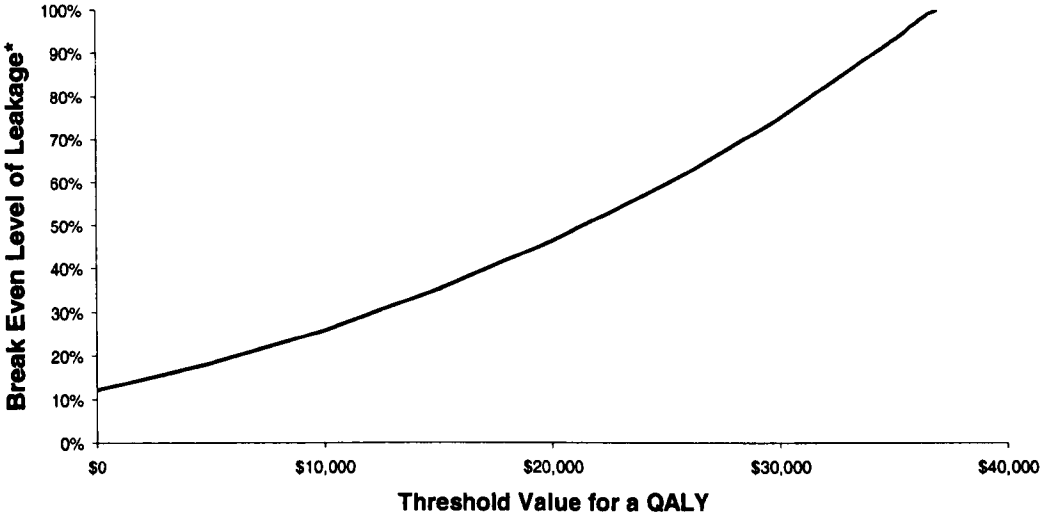
The health care costs associated with the optimal limited use criteria will increase as the value of λ increases as it becomes optimal to provide alendronate to more and more women. However, given the higher health care costs associated with the prescribing of etidronate, the costs associated with the use of optimal limited use criteria will be less than under current prescribing. Thus, a budget constraint based on current overall health expenditure on osteoporosis will not impact the optimal limited use criteria (Figure 7.8).

The drug costs associated with the optimal limited use criteria will increase as λ increases as it will become optimal to provide alendronate to more and more women.

---

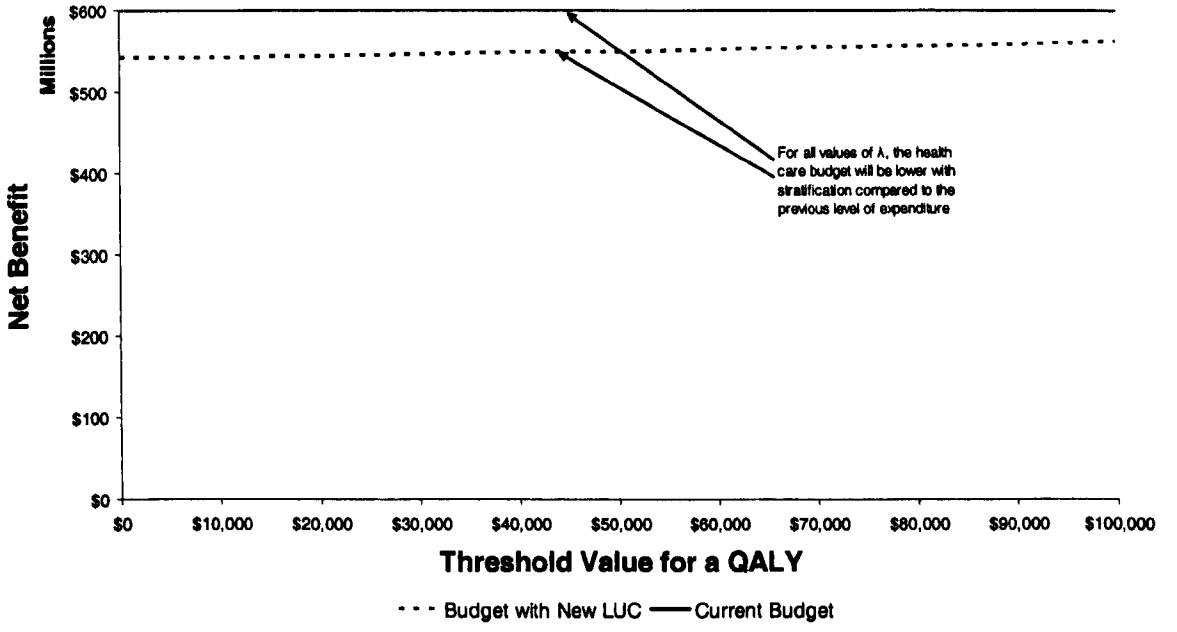
[68] $\Delta_i TNB = 0.46 * \$38 \text{ million} = \$17.46 \text{ million} = TNB_{sjk}$.

Figure 7.7: Break Even Level of Leakage Required for Switch to Alendronate Being Unavailable for all Women by Value of a QALY



* Level of leakage is percentage of patients who do not meet LUC who will receive treatment. Break even point is level of leakage where the net benefit loss from leakage is greater than the net benefit from stratification.

Figure 7.8: Health Care Budget with Current Prescribing and Revised Prescribing under New Limited Use Criteria by Value of a QALY



Given the higher drug costs of alendronate compared to etidronate, at λ greater than $39 000, the forecasted drug budget would be greater than the current drug budget (Figure 7.9).

If a budget constraint was introduced, then if λ was greater than $39 000, the optimal LUC would not be dependent only on whether the net benefit of alendronate was greater than zero. The focus of the LUC would be to maximize the return in the form of net benefit but within the prevailing budget constraint. For example, if λ was equal to $50 000, the optimal LUC would lead to a drug budget of $43.7 million. Thus, to remain within the budget constraint of $41.8 million a revised optimal LUC would be required. In this scenario the optimal LUC would be to restrict alendronate to women with previous fracture aged 75 and over and with no previous fracture aged 79 and over. Thus, the net benefit from stratification will be lower than without any budget constraint ($24.7 million compared to $24.9 million). This net benefit loss from the imposition of a budget constraint increases as λ increases as more and more patients for whom alendronate is optimal without a budget constraint are allowed access (Figure 7.10).

7.4.5 Comparison of Optimal Limited Use Criteria with Previous Limited Use Criteria

Figure 7.11 illustrates the potential net benefit gain if the Ontario Drug Benefit Formulary adopted a LUC based on the stratified analysis to replace the LUC in place at the beginning of 2003. The net benefit gain if λ was equal to $50 000 would be $93 million with the gain increasing the greater the value of λ.

Figure 7.9: Drug Budget with Current Prescribing and Revised Prescribing under New Limited Use Criteria by Value of a QALY

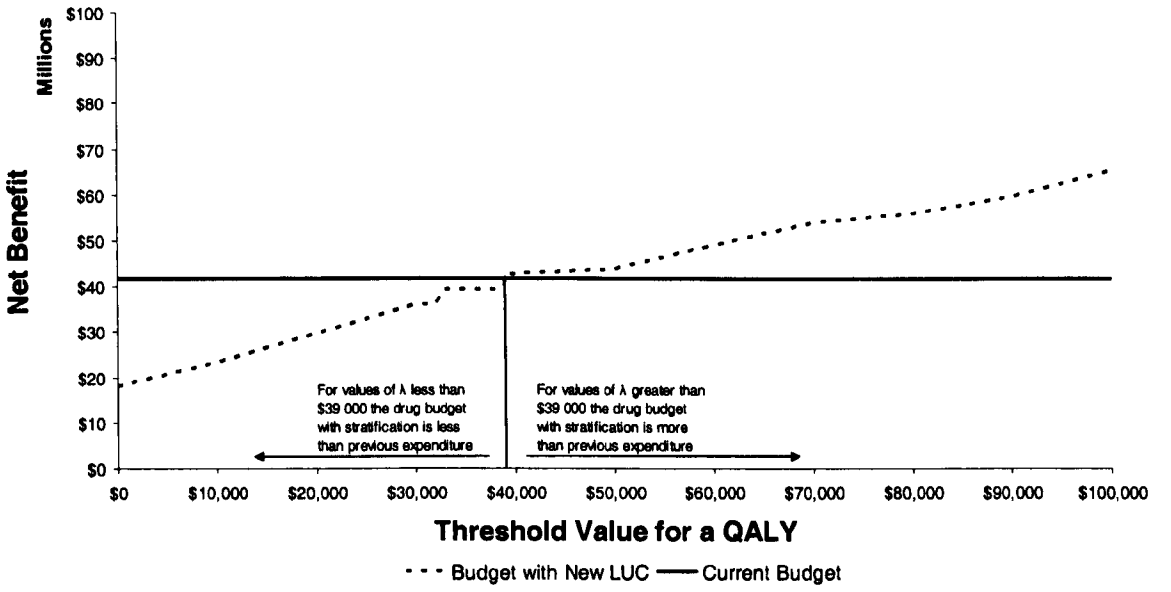


Figure 7.10: Net Benefit Loss Arising from Imposition of Constraint on Drug Budget

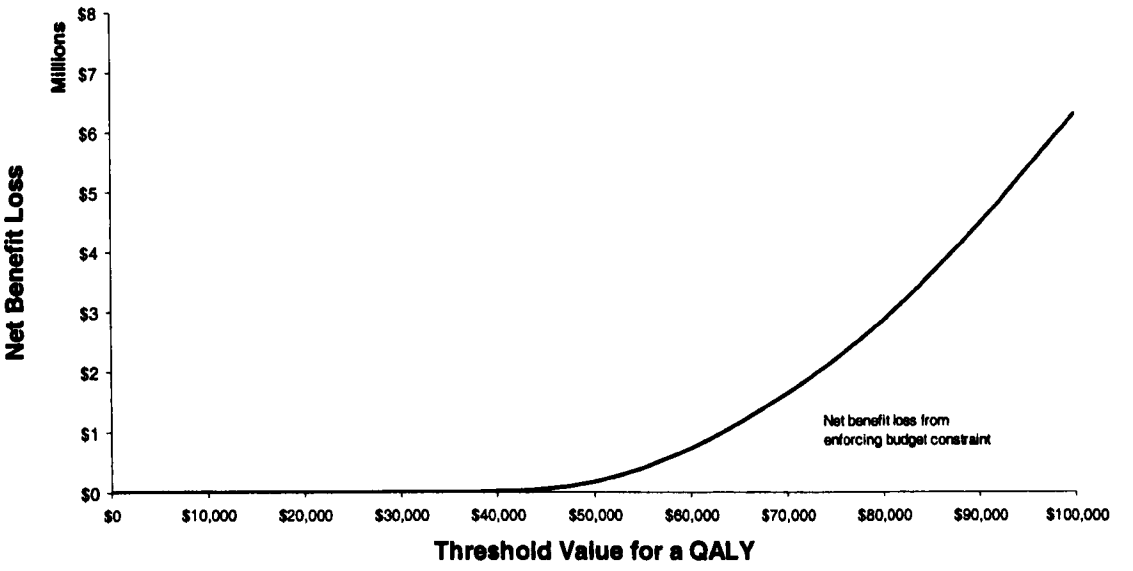
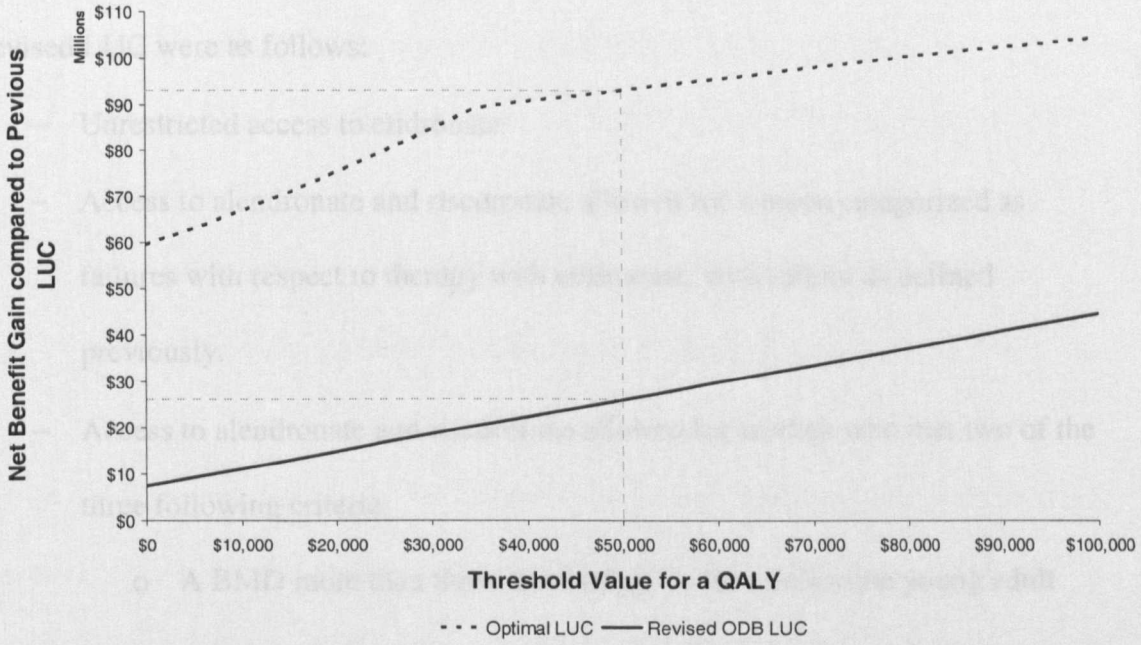


Figure 7.11: Net Benefit Gain from LUC Based on Stratified Analysis and Revised LUC from ODB Compared to Previous LUC



7.5 IMPACT OF ANALYSIS

Subsequent to an initial report of the above analysis (Coyle 2002), Ontario Drug Benefit Formulary revised the LUC for bisphosphonates in the spring of 2003[69]. The revised LUC were as follows:

- Unrestricted access to etidronate
- Access to alendronate and risedronate allowed for women categorized as failures with respect to therapy with etidronate; with failure as defined previously.
- Access to alendronate and risedronate allowed for women who met two of the three following criteria
    - A BMD more than three standard deviations below the young adult mean
    - A previous osteoporotic fracture
    - Aged 75 and above

The final criteria allows for increased access to alendronate and risedronate as alternate therapies to the use of etidronate. It is not possible to obtain an accurate measure of the net benefit gain from the adoption of the revised LUC given that it is unclear what the relative proportions of patients is who will meet the additional criteria. To explore this issue, the following assumptions relating to the potential market share of each bisphosphonate were made.

[69] A report prepared by me based on an abstract presented at the 2003 Medical Decision Making conference was part of a submission by Merck and Co. Canada seeking revised LUC for alendronate. Based on this submission, the LUC for bisphosphonates within the ODB formulary were revised.

- For all women who met the additional criteria for access to alendronate and risedronate, it was assumed that the market share for each bisphosphonate was equal.
- For all women who did not meet the additional criteria for access to alendronate and risedronate, it was assumed that the market share remained as previously.
- For cohorts it was assumed that 10% of women had a BMD less than 3 standard deviations below the young adult mean (Tenenhouse et al. 2000).

Clearly, the above assumptions can at best be considered tentative. Under these, the net benefit gain from the revised LUC was $26 million, if λ was equal to $50 000, with the gain increasing linearly the higher the value of λ (Figure 7.11). However, the revised LUC is inefficient when compared to the optimal LUC as defined by the stratified analysis, with a reduction in net benefit of $57 million when λ equals $50 000.

7.6 DISCUSSION

Previous to 2003, the use of bisphosphonates within the Ontario Drug Benefit plan was restricted through limited use criteria which led to etidronate comprising 74% of bisphosphonates covered. However, a meta analysis suggested there was limited if any benefit from the use of etidronate in terms of the prevention of fractures (Cranney et al. 2001a). Given this and additional evidence relating to alendronate and risedronate (Cranney et al. 2002a, 2002b), an economic analysis assessing the cost effectiveness of bisphosphonates became desirable.

A stratified analysis found that regardless of the value for λ, an optimal LUC will involve a proportion of women having access to alendronate with the rest having no access to therapy. Adoption of the optimal LUC as defined by the stratified analysis would lead to substantial health and economic benefits over the previous LUC. If λ was equal to $50 000, the net benefit gain from such a change would be $93 million.

Adoption of the LUC based on stratified analysis may be problematic for decision makers. This would require two major changes to policy relating to bisphosphonates. First, only one bisphosphonate would be made available for prescribing compared to the current three; albeit two are subject to restricted access. This may be considered a violation of physician autonomy; even though the therapy which is available is both the most effective and most expensive. Secondly, imposition of the LUC would result in a group of osteoporotic women having no access to therapy.

Given such concerns, it is maybe unsurprising that the revised Ontario Drug Benefit Formulary LUC are not wholly consistent with the results of this analysis. For all women, at least one bisphosphonate is made available for prescribing and women at high risk of fracture have unrestricted access to all three bisphosphonates. However, the revised LUC should lead to a move towards greater prescribing of alendronate and risedronate relative to etidronate.

The overall net benefit gain from the revised LUC will be unknown until data demonstrating the change in prescribing patterns are available. Tentative analysis suggests that noticeable gains in net benefit should be realized though these will be

substantially less than that which would be achieved from the LUC based strictly on the results of the stratified analysis.

7.7 CONCLUSION

In this chapter, the framework developed in Chapter 3 has been applied to the prescribing of bisphosphonates for the prevention of fractures in Canadian women. A stratified economic analysis was conducted using the Canadian Economic Model of Osteoporosis developed for this thesis and described in detail in Chapter 6.

Analysis estimated the lifetime costs and QALYs associated with three different therapies for the prevention of osteoporotic fractures as well as a strategy involving no therapy for 62 cohorts based on a woman's age and previous history of fracture. Analysis found that, an optimal LUC will involve a proportion of women having access to alendronate with the rest having no access to therapy. The proportion of women for which alendronate is optimal increases, the greater the value of λ, with alendronate use more cost effective for older women with previous fracture.

Further analysis examined the impact of both equity considerations and leakage on the net benefits to be gained from stratification. Analysis found that equity concerns relating to restricting access to therapy on the basis of age would have a significant effect on net benefit. Further analysis found that leakage could have substantial impact on net benefit and may lead to it being optimal to make all therapaies unavailable to patients.

The above analysis is based on the expected values of outcome generated through Monte Carlo simulation. This is the optimal method for determining an efficient allocation of resources. Consideration of the uncertainty over the optimal policy is appropriate for determining which future research is worthwhile. This is the focus for the following chapter.

Chapter 8.

Value of Information Analysis for the Treatment of Osteoporosis

---

8.1 INTRODUCTION

In Chapter 4, appropriate methods for handling uncertainty in input parameter values were identified. Uncertainty relates to lack of knowledge concerning a specific parameter. Given the potential for non linear relationships between input parameters and outcomes, probabilistic analysis based on Monte Carlo simulation is necessary to identify optimum therapies based on the expected values of costs and outcomes. There are opportunity losses arising from making decisions based on uncertain information which are a result of the potential for type 3 error. Thus, the principal focus of further analysis relating to uncertainty should be to determine for which parameters the collection of further information is justified.

In Chapter 4, the expected value of perfect partial information (EVPPI) was identified as the theoretically correct method for determining parameter importance and is a necessary component to the determination of an efficient research design. An approach for screening parameters was developed and methods for the estimation of optimal research design were described.

The focus of this chapter is to determine an efficient research plan relating to the treatment of osteoporosis. Analysis was conducted using a decision model for the

conduct of economic evaluations in osteoporosis for Canada described in Chapter 6. The results of the stratified analysis using this model were detailed in the previous chapter.

In Section 8.2, the methods adopted for the analysis are described and detailed in the form of the specific steps required. In Section 8.3, the optimal research plans for the potential research studies are identified. In Section 8.4, sensitivity analysis is conducted relating to the impact of uncertainty around features of a particular study design and the impact of alternate study designs.

8.2 METHODS FOR THE ANALYSIS OF INFORMATION VALUE

8.2.1 Steps involved in determining optimal sample size

Throughout Chapter 4, methods for estimating parameter importance and the value of further research are described and an approach to applying this to studies is developed. The approach can be outlined as a series of 5 steps:

1. Estimate EVPPI for individual parameters and parameter groups
2. Estimate population EVPPI for individual parameters and parameter groups
3. Estimate the costs associated with specific research projects
4. Eliminate studies where the minimal costs of research are greater than population EVPPI
5. For potential studies which are not eliminated estimate the optimal sample size.

Each of these steps is now discussed in detail.

8.2.2 Estimating EVPPI

In Chapter 4, 5 methods for identifying EVPPI were presented (Claxton et al. 2001, Brennan et al. 2002a, 2002b, Felli and Hazen 1998, Coyle et al. 2003c, Chilcott et al. 2003a). One method was dismissed as being an incorrect formulation of EVPPI (Claxton et al. 2001). Two methods were identified as being appropriate only in specific circumstances relating to the form of uncertainty around parameters and the relationship between incremental net benefit (Felli and Hazen 1998, Chilcott et al. 2003a). The two final methods were shown to be appropriate measures in all circumstances though both were computationally complex (Brennan et al. 2002a, 2002b, Coyle et al. 2003c).

Given the complexity and the number of uncertain parameters within the model, a screening mechanism is adopted as outlined in Chapter 4. A proxy value for EVPPI for all parameters and parameter groups are estimated using the single MCS method outlined in Section 4.4.3.2 (Felli and Hazen 1998). This method requires the following:

1. Conduct a MCS by sampling from the probability density functions of the parameters of interest (X_i) with all other parameters fixed at their expected value ($X^c = E(X^c)$).
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. A proxy for EVPPI is the expected value from step 2.

Within the osteoporosis model, there are 62 distinct analyses of cost effectiveness: one for each of 62 population strata. By definition, the EVPPI per patient for an individual parameter or group of parameters will be the sum of EVPPI for each of the 62 strata weighted by the proportion of pertinent patients in each strata.

$$EVPPI_{x_i} = \sum_{j=0,1,k=65,66..95} EVPPI_{x_i jk} * p_{jk}$$

where

p = *proportion of patients in each strata*

$j = 0$ *if no previous fracture*, 1 *if previous history of fracture*

k = *age* (*range from* 65 *to* 95)

EVPPI is estimated using the single MCS method for all 36 individual stochastic parameters in the model. In addition, EVPPI is estimated for the following 7 parameter sub-groups given the likelihood that data for each group could be collected within a single study design:

- The relative risk of hip, spine and wrist fracture for each therapy (3 groups)
- The cost of hip fracture by patient outcome and age
- Utility weights for hip, spine and wrist fractures
- Continuation rates for alendronate, etidronate and risedronate
- Age specific rates of admission to LTC following hip fracture

The list of parameter sub-groups is not exhaustive but is limited to the most likely research designs.

8.2.3 Estimating Population EVPPI

In the previous section the methods for estimating EVPPI are outlined. EVPPI is an estimate of the opportunity loss on a per patient basis. To estimate the value of

perfect information across all potential patients, EVPPI needs to be weighted by the potential patient population which will be affected by the research design.

The potential patient population is a function of three parameters: the number of patients who could be affected annually; the duration for which the research is meaningful and the likely uptake of research.

In this study, the number of future patients whose treatment could be affected as a result of further research is assumed equivalent to the number of incidence cases of prescribing of bisphosphonates. For the province of Ontario this is estimated to be 74 848 based on data for 2001. For the base analysis, this figure is assumed fixed given that it is based on actual population data.

However, it should be noted that there is uncertainty around the appropriateness of such an estimate. For certain parameters within the decision model, the value of information may be limited to only Ontario (e.g. costs of fractures). However, for some parameters, more precise estimates may have value for patients and decision makers elsewhere. For example, the results of a clinical trial may reduce uncertainty in decision making for Ontario but, in addition, will reduce uncertainty for many other jurisdictions. The appropriateness of including such factors on the decision over optimal research design will be determined by whether decision makers place any value on reducing decision uncertainty outside of their jurisdiction[70].

[70] For this chapter, analysis is restricted to assuming only benefit to patients within the province of Ontario. Clearly, including a wider basis for patients will increase the return from research leading to higher optimal sample sizes.

Consideration of the duration of time for which research can be considered meaningful (the lifespan of research) allows for the fact that the impact of research will become limited over time as different therapeutic options become available. For instance, refining our estimates of the effectiveness of etidronate will have no value once etidronate is no longer considered an appropriate option for treatment. For the base analysis it is assumed that the lifespan of research will be up to 10 years from the time of the current decision[71]. This estimate reflects the rate of licensing of new therapeutic options in this clinical area[72].

Many previous analyses of information value have failed to further adjust for the duration of time required for the research and the subsequent revision of any LUC. If a clinical trial takes up to 4 years to produce the desired information and leads to a policy change, then the potential lifespan of the research will be limited (in this example from 10 years to 6). For each potential research design, the time between commencement and a revision of LUC is estimated and where appropriate assumed to vary by sample size. The impact of shorter research duration (e.g. faster trial recruitment) is examined in Section 8.4.

The final factor which influences the population weighting is the uptake of any subsequent change in policy following the research. Under the scenario developed, an initial decision on the appropriate limited use criteria will be made based on current evidence. Subsequent evidence may lead to a revision of the criteria. If the

[71] For the base analysis, it is assumed that the lifespan of research is known and fixed. However, in reality this is unlikely. Thus, in section 8.4 the impact of assuming that variables are unknown is explored. For lifespan uncertainty is represented by a triangular distribution bounded by 7 and 13 years.

[72] For different clinical areas the lifespan of research may vary. For areas for which there is rapid progress in the development of therapeutic alternatives the lifespan of certain research may be quite short.

research does not or can not have the potential to lead to a revision there is by default no value to the information: i.e. the only value from information arises from the potential to revise the original decision.

The thesis adopts a normative framework for the decision makers who decide on which therapies should be reimbursed for which patients. However, as in the case of leakage discussed in Chapter 3, the thesis does not take the position of assuming that other decision makers such as individual clinical practitioners will be normative in their behaviour. Thus, not all practitioners may be willing to change their prescribing practice when revised criteria are put in place.

For the base analysis, it is assumed that the majority (75%) of practitioners will change according to any revised criteria[73]. This is assumed to be constant for all types of research but it is possible that compliance with new criteria will be a function of the nature of research which influenced any changes and the nature of any changes[74].

Given the above, the proxy value for population EVPPI based on the single MCS method can be estimated as follows:

[73] In section 8.4.1, uncertainty around this estimate is represented by a beta distribution with alpha =30 and beta = 10.

[74] Practitioners may be more likely to comply with changes based on new trial evidence which implies it is cost effective to treat a greater number of patients than evidence of falling costs in treating fractures which implies it is cost effective to treat a lesser number.

$$popEVPPI = EVPPI * u * n * \left(\sum_{y_{ls}=0...ls-1} \frac{1}{(1+r)^{y_{ls}}} - \sum_{y_d=0...d_{\min}-1} \frac{1}{(1+r)^{y_d}} \right)$$

where

u = *uptake of revised LUC by practitioners*

n = *annual number of potential patients affected by LUC*

ls = *lifespan of research*

$d_{\min}$ = *minimum duration of research*

8.2.4 Estimating the costs associated with specific research projects

In the previous section, the methods for estimating the population EVPPI for a specific parameter were described. Population EVPPI is the maximum benefit which could be obtained from research information. Thus, to determine the optimal sample size for a given research project, it is necessary to provide an estimate of the associated costs of the project.

Costs will vary by study sample size. However for any research project there will be fixed costs associated with start up costs relating to the study design and implementation stages and finishing costs relating to the analysis and presentation of study data.

In Canada, research budgets funded through the Canadian Institutes of Heath Research and other peer reviewed funding agencies are underestimates of the true cost of conducting research. Budgets exclude both the costs of developing research and the substantial time commitments of study investigators and co investigators which are not covered by funders. For this analysis, time commitments are assumed to be quantifiable as a fixed annual cost. In addition it may be necessary to include

the costs of revising the economic analysis with the new information and updating policy with respect to the treatments evaluated.

The research costs will be partially determined by the rate by which participants can be recruited into a study as well as the time required to start up a study. The slower the recruitment to the study, the less the discounted cost per case and the lower the discounted finishing costs. However, this is likely to be outweighed by the higher investigator costs due to the longer study duration. Similarly, the longer the duration of time required to start up the study, the greater the cost of investigator's time.

Thus, for each potential research study[75], the following parameters are required to allow estimation of research costs:

- start up costs (c_{su})
- time required for start up (d_{su})
- costs per sample (c_{case})
- finishing costs (c_{fin})
- time required for finishing (d_{fin})
- analysis and policy costs (c_{ap})
- time required for analysis and policy revision (d_{ap})
- annual recruitment rates (ARR)
- duration of follow up of patients (d_{flup})
- annual costs relating to research administration and investigator time (c_{annual})
- discount rate (r)

[75] Potential research studies are limited by focusing on those with non negligible EVPPI.

The total research costs (TC) as a function of the sample size (s) can be estimated as follows:

$$TC_s = c_{su} + \sum_{y=0,\ldots d-1} \frac{c_{annual}}{(1+r)^y} + ARR * c_{case} * (1 - d_{su}) + \sum_{y=1,\ldots d-1-d_{fin}} \frac{c_{case} * ARR}{(1+r)^y} + \frac{(c_{fin} + c_{ap})}{(1+r)^{d-1}}$$

where

$$d = \frac{s}{ARR} + d_{flup} + d_{su} + d_{fin} + d_{ap}$$

8.2.5 Eliminating studies with no information value

For each potential research project, a minimum sample size (s_{min}) can be estimated which reflects the minimum for which it will be practical to conduct the research project. In the previous section, methods for estimating the costs of research projects with different sample sizes were detailed. Thus, the total cost for the minimum sample size can be estimated as:

$$TC_{s_{min}} = c_{su} + \sum_{y=0,\ldots d_{min}-1} \frac{c_{annual}}{(1+r)^y} + ARR * c_{case} * (1 - d_{su}) + \sum_{y=1,\ldots d_{min}-1} \frac{c_{case} * ARR}{(1+r)^y} + \frac{(c_{fin} + c_{ap})}{(1+r)^{d_{min}-1}}$$

where

$$d_{min} = \frac{s_{min}}{ARR} + d_{flup} + d_{su} + d_{fin} + d_{ap}$$

The population EVPPI which is estimated as in Section 8.2.3 can be seen as the maximum value that can be obtained from research. Thus, if the population EVPPI is less than the minimum cost of research a proposed research project can not have any information value. Thus, it is desirable to eliminate potential research projects where the maximum net benefit from sampling (NBS_{max}) is negative:

$$NBS_{max} = popEVPPI - TC_{s_{min}}$$

8.2.6 Estimating optimal sample size for studies

The previous four sections detail a method for reducing the number of potential study designs for which it is necessary to conduct value of information analysis. Once these steps are conducted a reduced number of study designs will be considered and the optimal sample size for each study design identified based on the methods outlined in Chapter 4.

Estimation of optimal sample size for each study involves the following steps: EVPPI is estimated for relevant parameters based on current data using the quadrature method. This involves the following:

1. A set of 101 values is determined for each parameter of interest. The values are equally spaced across the parameters probability function with a high degree of coverage (>99.99%).
2. For each value of the parameter chosen in step 1, the NB for all treatment options is estimated by conducting MCS by sampling from the probability density functions of all other parameters (X^c).
3. For each simulation conducted in step 2, the net benefit of the optimum therapy from the base analysis is subtracted from the maximum net benefit over all therapeutic options.
4. Each estimate from step 3 is weighted by the probability density for the specific value of the parameter.
5. EVPPI is then estimated by numerically integrating across the probability density function.

Data collection from a hypothetical study with sample size s is simulated based on current knowledge. This requires the following:

1. Draw a sample from the prior distribution of each of the parameters of interest.
2. Use this sample to create a new distribution which relates to a single participant in the potential study.
3. Sample from this distribution[76].
4. The probability distribution(s) for input parameter(s) is updated by combining prior (original) data with simulated sample data.
5. EVPPI|*s* is estimated based on the updated distribution. This requires merely re-estimating EVPPI as in step 1 by changing the probability density for parameter values as outlined in step 4.
6. Steps 1 – 4 are repeated a number of times (3000) to allow estimation of the expected value of EVPPI|s

$$EVPPI|s = \frac{\sum EVPPI|s_r}{3000}$$

7. The population expected value of sample information (popEVSI) is the result of the subtraction of the value from step 1 and the value from step 6 weighted by the number of future patients potentially affected by the research

$$popEVSI_s = (EVPPI - EVPPI|s) * u * n * \left(\sum_{y_{ls}=0...ls-1} \frac{1}{(1+r)^{y_{ls}}} - \sum_{y_d=0...d_s-1} \frac{1}{(1+r)^{y_d}} \right)$$

where
u = uptake of revised LUC by practitioners
n = annual number of potential patients affected by LUC
ls = lifespan of research
d_s *= duration of research given sample size s*

---

[76] For example for a beta distribution the first step involves randomly sampling a probability value from the beta distribution. The second step involves deriving a binomial distribution with the probability value as sampled above and with s number of trials. The third step involves simulating a study with sample size s from this binomial distribution.

8. Steps 1 – 6 are repeated for various s.

The optimal sample size (s*) is the sample size which maximizes the difference between the EVSI and the cost of sample information

$$s^* = max_s \left(popEVSI_s - TC_s \right)$$

8.3 RESULTS

8.3.1 EVPPI for Individual Parameters and Parameter Sub-groups

Table 8.1 presents the proxy estimates for EVPPI all individual parameters with non zero EVPPI and for the 7 identified sub-groups. The following are the major findings:

- The highest EVPPI relates to obtaining further information on the effectiveness of etidronate in preventing hip fracture ($215.92); followed by the same information for alendronate ($19.63).
- For both etidronate and alendronate, there is little additional value in obtaining further information on the prevention of wrist and spine fractures highlighting that further research if conducted could focus on only collecting information concerning hip fractures.
- The EVPPI for further information relating to risedronate is minimal ($0.05).
- The EVPPI for individual cost parameters are zero whilst the EVPPI for all cost parameters relating to hip fractures is small ($0.08).
- The EVPPI for utilities is similar to the EVPPI for the utility of hip fracture ($0.08). This suggests that there may only be only value in refining the latter variable.

Table 8.1: EVPPI for Individual Parameters and Parameter Sub-groups

Individual parameter	EVPPI per patient
RR of hip fracture with etidronate compared to no therapy	$215.92
RR of hip fracture with alendronate compared to no therapy	$19.63
Proportional reduction in benefit after stopping therapy	$5.02
RR of fracture with previous fracture compared to osteoporosis without fracture	$1.56
Proportion of LTC stay post hip fracture attributable to hip fracture	$1.19
Rate of admission to LTC post hip fracture for women aged 75-84	$0.28
RR of hip fracture with osteoporosis compared to no osteoporosis	$0.19
Disutility associated with hip fracture	$0.08
RR of hip fracture with risedronate compared to no therapy	$0.05
RR of spine fracture with etidronate compared to no therapy	$0.05
RR of spine fracture with alendronate compared to no therapy	$0.05
RR of spine fracture with risedronate compared to no therapy	$0.05
Parameter group	
RR of hip, wrist and spine fractures with etidronate compared to no therapy	$216.14
RR of hip, wrist and spine fractures with alendronate compared to no therapy	$19.65
Rate of admission to LTC post hip fracture for women for all age groups	$0.28
Disutility associated with hip, wrist and spine fracture	$0.08
Costs of hip fractures by age and outcome	$0.05
RR of hip, wrist and spine fractures with risedronate compared to no therapy	$0.05
Annual continuation rates for etidronate, alendronate and risedronate	$0.05

EVPPI for all individual parameters not listed equals $0.
RR = relative risk

- The EVPPI for the rates of institutionalization post hip fracture is similar to the EVPPI for just the rate for women aged 75-84 ($0.28). This again suggests that there may only be only value in refining the latter variable.
- Individually, continuation rates have no further information value whilst collectively the EVPPI is low ($0.05).
- There may be value in obtaining further information relating to the relative risk of fracture with previous history (EVPPI = $1.56), the benefit beyond curtailment of therapy (EVPPI=$5.02) and the proportion of LTC attributable to fractures (EVPPI=$1.19).

Given these findings it is necessary to evaluate the value of information relating to the following parameters.

- Relative risk of hip fracture on etidronate
- Relative risk of hip fracture on alendronate
- Relative continuation of benefit after stopping therapy
- Relative risk of having a fracture associated with a previous fracture
- Proportion of long term care stays attributable to hip fracture
- Rate of admission to long term care after hip fracture for women aged 75-84
- Relative risk of fracture associated with osteoporosis
- Disutility associated with hip fracture
- Annual continuation rates with therapy for alendronate, etidronate and risedronate
- Costs associated with hip fractures by age and patient outcome.

8.3.2 Population EVPPI for Individual Parameters and Parameter Groups

In Section 8.2.2, parameters relating to the number of potential patients for whom research may affect their future care were identified. The annual number of patients affected was 74848, the lifespan of research was 10 years and the uptake of any new LUC was 75%. Thus, the total number of potential patients is 561 300 undiscounted and 455 141 undiscounted.

To determine the population EVPPI for each parameter(s) identified in the previous section it is necessary to revise the total number of potential patients by allowing for the minimum time for which research can be conducted and change policy. For each pertinent parameter and parameter group the following information was estimated

- minimum sample size (s_{min})
- time required for start up (d_{su})
- time required for finishing (d_{fin})
- time required for analysis and policy revision (d_{ap})
- annual recruitment rates (ARR)
- duration of follow up of patients (d_{flup})

The minimum duration of research for each of the 10 potential studies and the study design are presented in Table 8.2. Potential research projects can be characterised as clinical trials (for relative risks associated with treatments), case control studies (for relative risks associated with risk factors), cohort studies (for utilities, costs and transition probabilities) and analysis of administrative data (for continuation rates). Clinical trials are likely by their nature to have the greatest study duration whilst analysis of existing data will have the lowest.

Table 8.2: Assumed Features of Potential Study Designs

	Study Design	s_{min}	d_{su}	d_{fin}	d_{ap}	ARR	d_{flup}	d_{min}
RR of hip fracture with etidronate compared to no therapy	Clinical trial	200	0.25	0.17	0.17	365	3.00	4.13
RR of hip fracture with alendronate compared to no therapy	Clinical trial	200	0.25	0.17	0.17	365	3.00	4.13
Proportional reduction in benefit after stopping therapy	Cohort	200	0.25	0.17	0.17	365	1.00	2.13
RR of fracture with previous fracture compared to osteoporosis without fracture		200	0.25	0.17	0.17	365	1.00	2.13
Proportion of LTC stay post hip fracture attributable to hip fracture	Cohort	20	0.17	0.17	0.17	365	0.50	1.05
Rate of admission to LTC post hip fracture for women aged 75-84	Cohort	20	0.17	0.08	0.17	365	0.08	0.55
RR of hip fracture with osteoporosis compared to no osteoporosis	Case-control	200	0.25	0.17	0.17	365	1.00	2.13
Disutility associated with hip fracture	Utility exercise	20	0.17	0.08	0.17	365	0.08	0.55
First year costs of hip fractures by age and outcome	Costing study	20	0.17	0.08	0.17	365	1.00	1.47
Annual continuation rates for etidronate, alendronate and risedronate	Administrative data analysis	0	0.08	0.08	0.17	0	0.00	0.33

s_{min} minimum sample size
d_{su} time required for start up
d_{fin} time required for finishing
d_{ap} time required for analysis and policy revision
ARR annual recruitment rates
d_{flup} duration of follow up of patients
d_{min} minimum duration of study

The different minimum durations will lead to different potential patient populations with regards to the impact of the research studies. By weighting the EVPPI per patient by the potential patient population we can estimate the maximum value for each research study (Table 8.3). The findings of this are similar to that of EVPPI per patient in that the highest value is for the conduct of clinical trials of etidronate ($51.8 million) and alendronate ($4.7 million) with little value for studies of hip fracture costs ($19 000) and utilities ($34 000) and for continuation rates($22 000).

8.3.3 Estimating the costs associated with specific research projects

The next step in evaluating the value of further information was to determine the costs associated with specific research projects related to each parameter. The minimum costs for the research project must be less than the population EVPPI for the particular parameter(s) for there to be potential for information value.

The minimum costs for each specific research project was calculated based on the data described in the previous section plus the following data relating to the costs of conducting and using research:

- start up costs (c_{su})
- costs per sample (c_{case})
- finishing costs (c_{fin})
- analysis and policy costs (c_{ap})
- annual costs relating to research administration and investigator time (c_{annual})

Table 8.3 Population EVPPI for Potential Study Designs

	EVPPI per patient	Potential Patient Population	Population EVPPI*
RR of hip fracture with etidronate compared to no therapy	$215.92	239940	$51,807,911
RR of hip fracture with alendronate compared to no therapy	$19.63	239940	$4,710,028
Proportional reduction in benefit after stopping therapy	$5.02	338715	$1,700,349
RR of fracture with previous fracture compared to osteoporosis without fracture	$1.56	338715	$528,395
Proportion of LTC stay post hip fracture attributable to hip fracture	$1.19	396007	$471,249
Rate of admission to LTC post hip fracture for women aged 75-84	$0.28	423997	$118,719
RR of hip fracture with osteoporosis compared to no osteoporosis	$0.19	338715	$64,356
Disutility associated with hip fracture	$0.08	423997	$33,920
Costs of hip fractures by age and outcome	$0.05	373474	$18,674
Annual continuation rates for etidronate, alendronate and risedronate	$0.05	436429	$21,821

* Discounted at 5% per annum

Note: Potential patient population varies due to the differences in time to the availability of evidence.

The above parameters were estimated based on discussion with other researchers and through using prior experience being a member of the Canadian Institutes of Health Information Health Services and Evaluation Committee (Table 8.4).

8.3.4 Eliminating studies with no information value

The final step before assessing the optimal sample size for potential studies is to eliminate any study which has no information value: i.e. the minimum costs of research are greater than the maximum benefit (population EVPPI).

Table 8.5 details the maximum net benefit from sample information for the ten potential studies. Six studies have a positive maximum net benefit and are subject to analysis to identify optimal sample size:

- RR of hip fracture with etidronate compared to no therapy
- RR of hip fracture with alendronate compared to no therapy
- Proportional reduction in benefit after stopping therapy
- RR of fracture with previous fracture compared to osteoporosis without fracture
- Proportion of LTC stay post hip fracture attributable to hip fracture
- Rate of admission to LTC post hip fracture for women aged 75-84

For the other four studies, the minimum cost of research exceeds population EVPPI signifying there is no value in further research:

- RR of hip fracture with osteoporosis compared to no osteoporosis
- Disutility associated with hip fracture
- First year costs of hip fractures by age and outcome

Table 8.4 Costs of Potential Study Designs

	c_{su}	c_{annual}	c_{case}	c_{fin}	c_{ap}	Minimum Cost of Research
RR of hip fracture with etidronate compared to no therapy	$30 000	$50 000	$500	$10 000	$10 000	$331 219
RR of hip fracture with alendronate compared to no therapy	$30 000	$50 000	$500	$10 000	$10 000	$331 219
Proportional reduction in benefit after stopping therapy	$30 000	$30 000	$300	$10 000	$10 000	$165 830
RR of fracture with previous fracture compared to osteoporosis without fracture	$30 000	$50 000	$300	$10 000	$10 000	$204 219
Proportion of LTC stay post hip fracture attributable to hip fracture	$20 000	$25 000	$500	$5 000	$10 000	$66 606
Rate of admission to LTC post hip fracture for women aged 75-84	$20 000	$25 000	$50	$5 000	$10 000	$45 442
RR of hip fracture with osteoporosis compared to no osteoporosis	$30 000	$50 000	$300	$10 000	$10 000	$204 219
Disutility associated with hip fracture	$10 000	$25 000	$50	$5 000	$10 000	$35 442
Costs of hip fractures by age and outcome	$10 000	$25 000	$50	$5 000	$10 000	$57 474
Annual continuation rates for etidronate, alendronate and risedronate	$10 000	$1 000	0	$5 000	$10 000	$25 065

c_{su} start up costs
c_{annual} annual costs of study management
c_{case} costs per sample
c_{fin} finishing costs
c_{ap} analysis and policy costs

Table 8.5: Maximum Possible Net Benefit from Potential Study Designs

	Population EVPPI	Minimum Cost of Research	Maximum Possible Net Benefit of Sample Information	Possibility that Research Study may be Worthwhile
RR of hip fracture with etidronate compared to no therapy	$51 807 911	$331 219	$51 476 691	Yes
RR of hip fracture with alendronate compared to no therapy	$4 710 028	$331 219	$4 378 809	Yes
Proportional reduction in benefit after stopping therapy	$1 700 349	$165 830	$1 534 519	Yes
RR of fracture with previous fracture compared to osteoporosis without fracture	$528 395	$204 219	$324 176	Yes
Proportion of LTC stay post hip fracture attributable to hip fracture	$471 249	$66 606	$404 642	Yes
Rate of admission to LTC post hip fracture for women aged 75-84	$118 719	$45 442	$73 277	Yes
RR of hip fracture with osteoporosis compared to no osteoporosis	$64 356	$204 219	-$139 864	No
Disutility associated with hip fracture	$33 920	$35 442	-$1 522	No
Costs of hip fractures by age and outcome	$18 674	$57 474	-$38 801	No
Annual continuation rates for etidronate, alendronate and risedronate	$21 821	$25 065	-$3 244	No

- Annual continuation rates for etidronate, alendronate and risedronate

These four studies have the lowest population EVPPI. For the RR of hip fracture with osteoporosis compared to no osteoporosis, this is primarily because of the relative lack of uncertainty over this parameter. For others, this is a result of either the low incidence of hip fractures (costs and disutilities) or the relatively low correlation between the parameter and the estimate of net benefit (continuation rates).

8.3.5 Estimating optimal sample size for studies

8.3.5.1 Relative risk of hip fracture with etidronate compared to no therapy

Further information relating to the relative risk of hip fracture with etidronate is assumed to come from a randomized controlled trial comparing etidronate to no therapy. Value of information analysis was conducted to assess the optimal sample size for such a trial with the objective of maximizing the net benefit of sample information.

To simulate the conduct of the trial and its results, the methods outlined in section 8.2.6 were used. The annual risk of hip fracture without treatment was assumed to be that of the control group in the FIT trial; 0.63% (Beta (α=40, β=6285)) (Black et al. 1996, Cummings et al. 1998). The relative risk of hip fracture with etidronate was that assumed in the base model (μ=0.945, σ=2.32).

Thus, to simulate the trial, a patient receiving no therapy is assumed to have an annual risk of hip fracture depicted by a binomial distribution with risk of event

equivalent to a sample value from the beta distribution. For a patient receiving etidronate the risk of hip fracture is characterized by a binomial distribution with risk of event equivalent to the sample value from the beta distribution weighted by the sample value from the lognormal distribution.

Figure 8.1 depicts the cost of sample information and EVSI by potential sample size for the clinical trial. EVSI is shown to rise significantly with information from the first few patients within a trial. For sample sizes greater than 650 EVSI is shown to fall as the value obtained from additional patients is less than the net benefit loss arising from the delay in obtaining information. The optimal sample size for the clinical trial is 640 patients with a net benefit of sample information of $43.6 million (Table 8.6).

8.3.5.2 Relative risk of hip fracture with alendronate compared to no therapy

Value of information concerning the relative risk of hip fracture with alendronate was assessed adopting the same methods as for etidronate above. The baseline risk of hip fracture is again assumed to be that of the control group in the FIT trial and the relative risk of hip fracture with alendronate is that assumed in the base model (μ=0.465, σ=1.29).

Figure 8.2 depicts the cost of sample information and EVSI by potential sample size for the clinical trial. EVSI is shown to rise almost linearly by sample size up to a sample of 800 patients. For sample sizes less than 220, a trial would not be worthwhile as the value of information will be less than the costs of the trial given the fixed start up and finishing costs. For sample sizes greater than 950 EVSI is

Figure 8.1 EVSI and Cost of Sample Information for Etidronate Trial by Sample Size

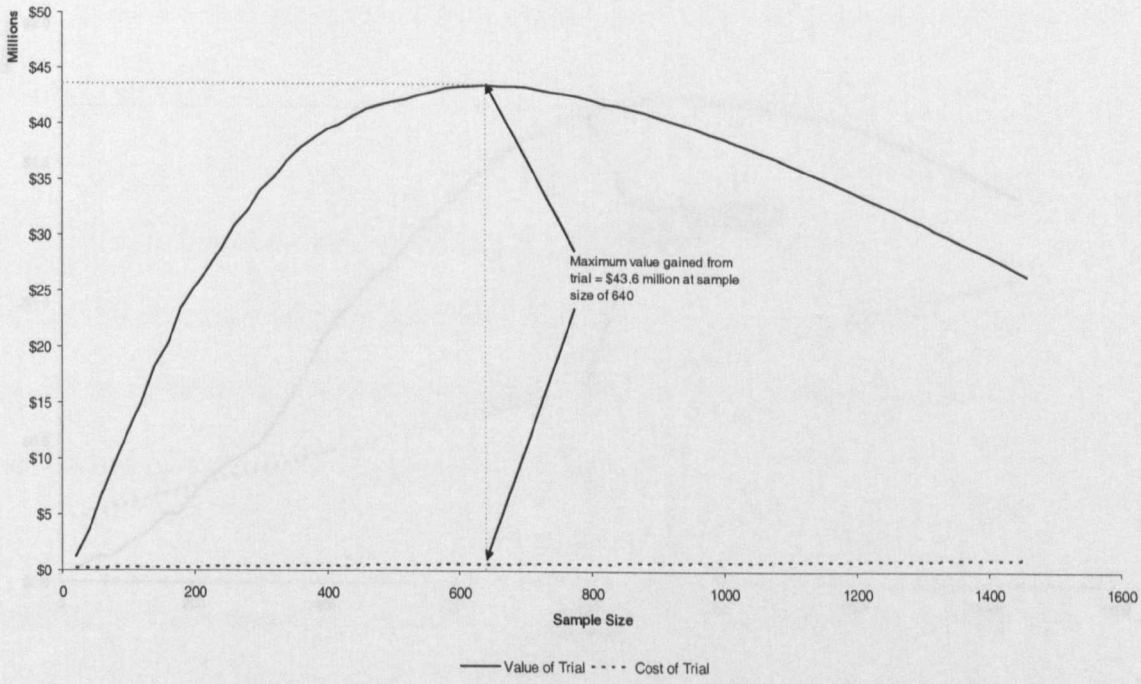
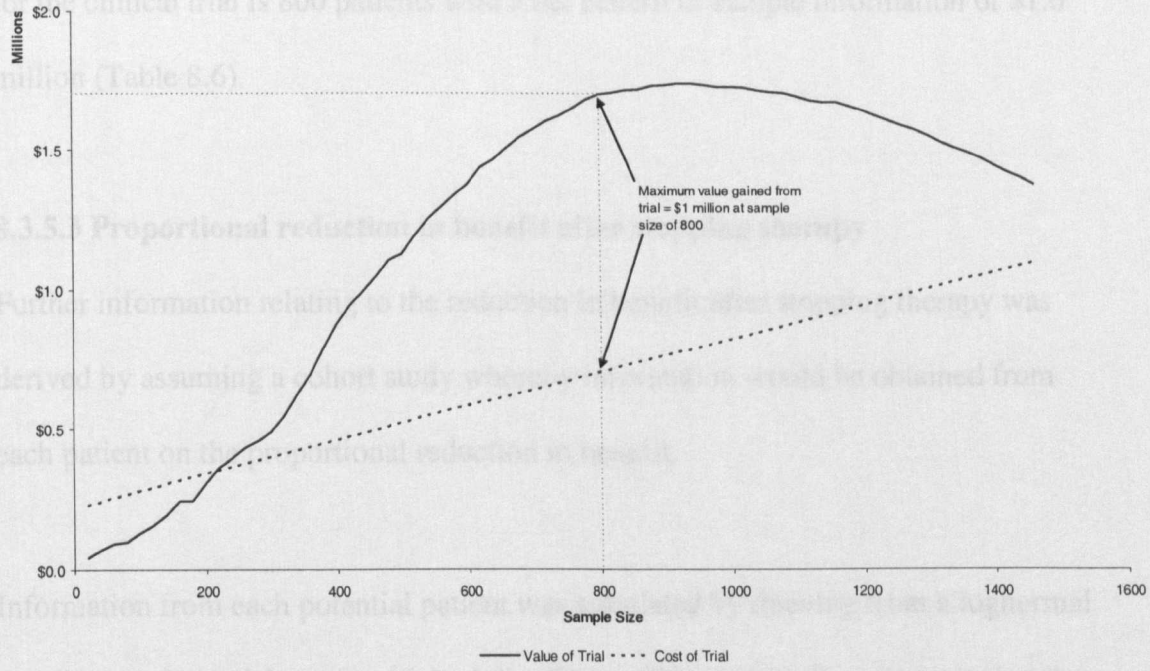


Table 8.6: Optimal Sample Size and Maximum Net Benefit of Sample Information

Parameter(s)	Study Design	Optimal Sample Size	Net Benefit from Sample Information
RR of hip fracture with etidronate compared to no therapy	Clinical trial	640	$43.6 million
RR of hip fracture with alendronate compared to no therapy	Clinical trial	800	$1.0 million
Proportional reduction in benefit after stopping therapy	Cohort study	350	$150 000
RR of fracture with previous fracture compared to osteoporosis without fracture	Case control	0	0
Proportion of LTC stay post hip fracture attributable to hip fracture	Cohort study	40	$510 000
Rate of admission to LTC post hip fracture for women aged 75-84	Cohort study	0	0
RR of hip fracture with etidronate compared to no therapy and with alendronate compared to no therapy	Clinical trial	750	$37.1 million

Figure 8.2 EVSI and Cost of Sample Information for Alendronate Trial by Sample Size



shown to fall as the value obtained from additional patients is less than the net benefit loss arising from the delay in obtaining information. The optimal sample size for the clinical trial is 800 patients with a net benefit of sample information of $1.0 million (Table 8.6).

8.3.5.3 Proportional reduction in benefit after stopping therapy

Further information relating to the reduction in benefit after stopping therapy was derived by assuming a cohort study whereby information would be obtained from each patient on the proportional reduction in benefit.

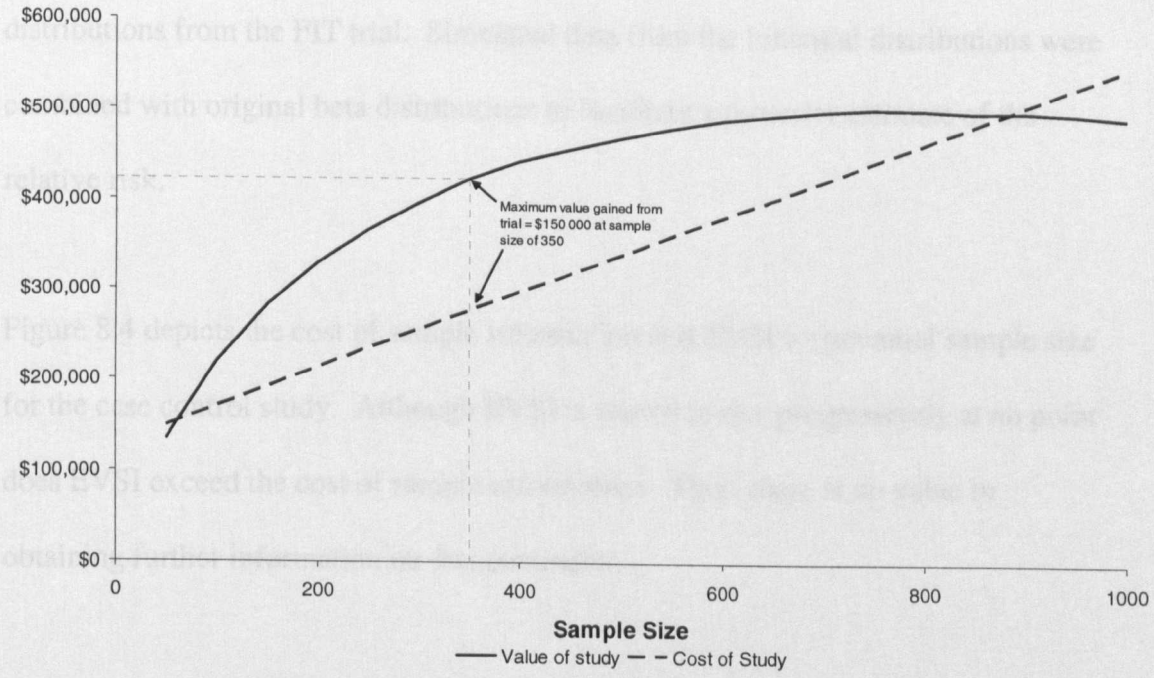
Information from each potential patient was simulated by drawing from a lognormal distribution derived from the original distribution (Normal (μ=0, σ=1) assuming the initial sample size was uncertain (uniform distribution between 0 and 2000).

Figure 8.3 depicts the cost of sample information and EVSI by potential sample size for the cohort study. EVSI is shown to rise progressively but decline for sample sizes greater than 750. The optimal sample size for the cohort study is 350 with a net benefit of sample information of $150 000 (Table 8.6).

8.3.5.4 Relative risk of fracture with previous fracture compared to osteoporosis without fracture

It was assumed that further information relating to the relative risk of patients having a fracture given a previous fracture compared to patients with osteoporosis without previous fracture could be obtained from a case control study.

Figure 8.3 EVSI and Cost of Sample Information for Cohort Study Examining Benefit beyond Treatment by Sample Size



Data for the case control study was simulated from data from the FIT trial which was the original source of the relative risk data. Binomial distributions for patients with and without previous fracture were obtained by drawing from the original beta distributions from the FIT trial. Simulated data from the binomial distributions were combined with original beta distributions to facilitate a posterior estimate of the relative risk.

Figure 8.4 depicts the cost of sample information and EVSI by potential sample size for the case control study. Although EVSI is shown to rise progressively at no point does EVSI exceed the cost of sample information. Thus, there is no value in obtaining further information on this parameter.

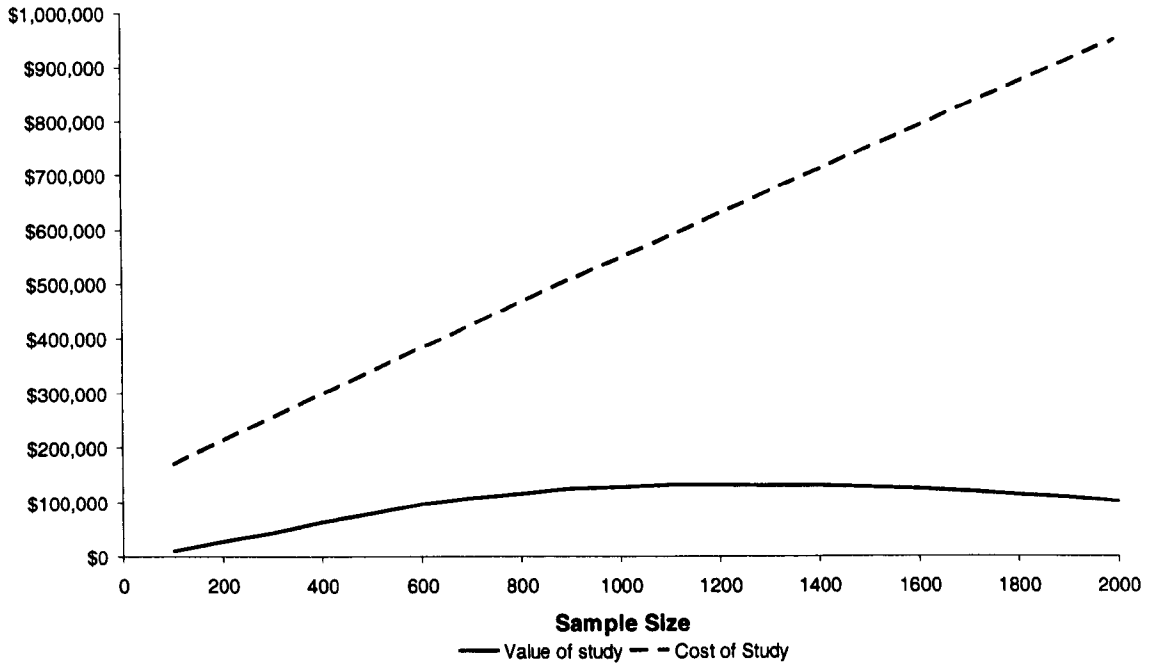
8.3.5.5 Proportion of LTC stay post hip fracture attributable to hip fracture

The prior estimate for the proportion of long term care stay attributable to hip fracture was based on limited information and was assumed to be uninformative (Beta (α=1, β=1)).

Further information on this parameter was assumed to come from a cohort study where estimates of the parameter for each subject were obtained from drawing from a binomial distribution with the expected value obtained from drawing from the original beta distribution.

Given the uncertainty concerning this parameter it is not surprising that there is substantial value to be obtained from the first samples within a cohort study but the

Figure 8.4 EVSI and Cost of Sample Information for Case Control Study of the Relative Increase in Fracture Risk given Previous Fracture by Sample Size



marginal value of sample information declines quickly (Figure 8.5). The optimal sample size for the cohort study is 40 with a net benefit of $510 000 (Table 8.6).

8.3.5.6 Rate of admission to LTC post hip fracture for women aged 75-84

The prior estimate for admission to LTC post hip fracture for women aged 75-84 was obtained from a cohort study (Beta (α=26, β=131)).

Further information on this parameter was assumed to come from a cohort study where estimates of the parameter for each subject were obtained from drawing from a binomial distribution with the expected value obtained from drawing from the original beta distribution. The posterior distribution is characterized by the sum of the number of women who were admitted to long term care in both cohort studies, and the sum of the number of women who were not admitted.

Figure 8.6 depicts the cost of sample information and EVSI by potential sample size for the proposed study. Although EVSI is shown to rise progressively at no point does EVSI exceed the cost of sample information. Thus, there is no value in obtaining further information on this parameter.

Figure 8.5 EVSI and Cost of Sample Information for Cohort Study Assessing the Proportion of LTC Stay Attributable to Fracture by Sample Size

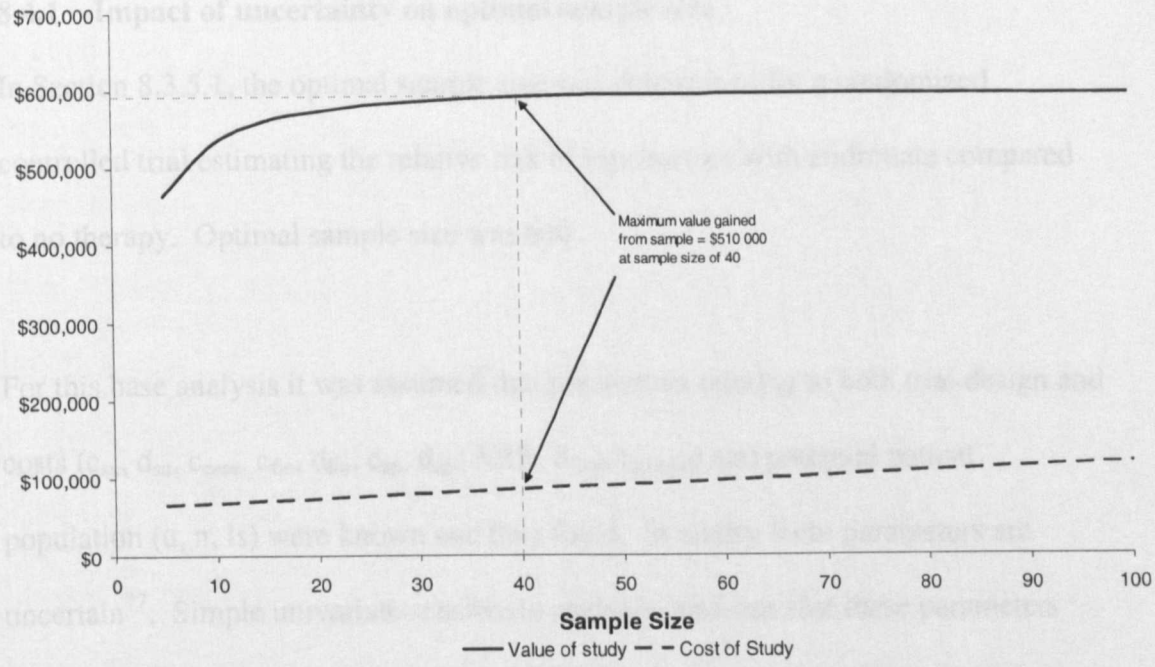
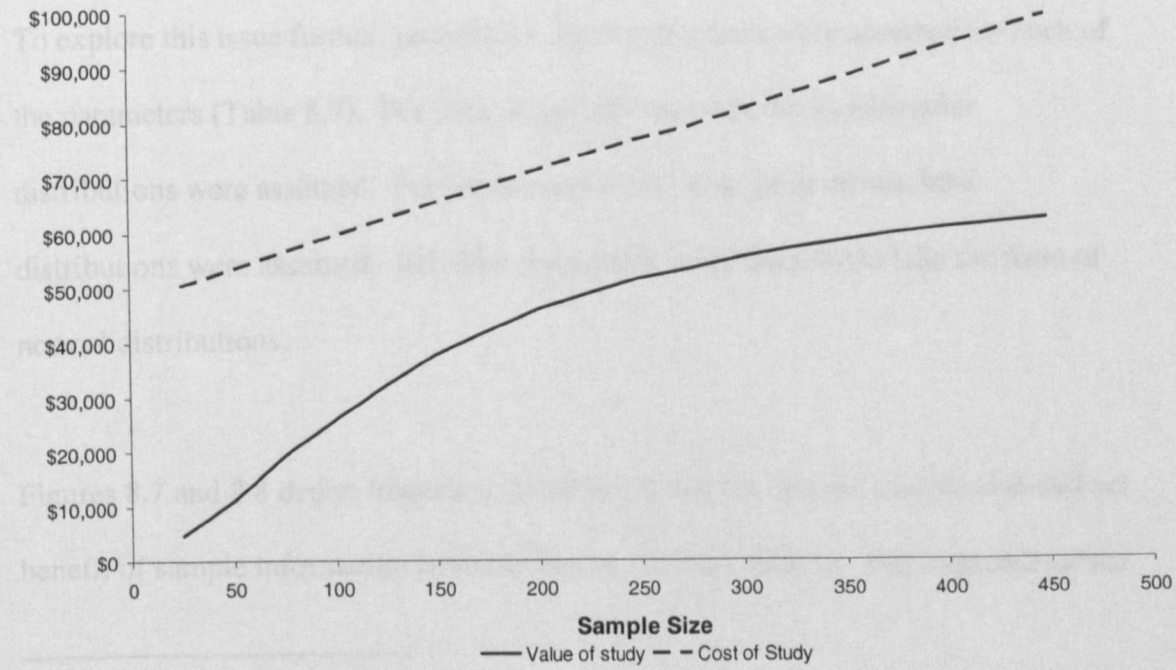


Figure 8.6 EVSI and Cost of Sample Information for Cohort Study Assessing the Proportion of Hip Fracture Patient aged 75-84 Admitted to LTC by Sample Size



8.4 FURTHER ANALYSIS

8.4.1 Impact of uncertainty on optimal sample size

In Section 8.3.5.1, the optimal sample size was determined for a randomized controlled trial estimating the relative risk of hip fracture with etidronate compared to no therapy. Optimal sample size was 640.

For this base analysis it was assumed that parameters relating to both trial design and costs (c_{su}, d_{su}, c_{case}, c_{fin}, d_{fin}, c_{ap}, d_{ap}, ARR, d_{flup}, c_{annual}) and potential patient population (u, n, ls) were known and thus fixed. In reality these parameters are uncertain[77]. Simple univariate sensitivity analysis confirms that these parameters have a non linear relationship with the optimal sample size. Thus, there is the potential that the expected value of the optimal sample size given the uncertainty around these parameters will be different than the optimal sample size based on the parameter's expected value.

To explore this issue further, probability density functions were assumed for each of the parameters (Table 8.7). For parameters relating to duration, triangular distributions were assumed. For parameters which were proportions, beta distributions were assumed. All other parameters were assumed to take the form of normal distributions.

Figures 8.7 and 8.8 depict frequency distributions for the optimal sample size and net benefit of sample information from the Monte Carlo simulation. The expected values

[77] Except for the duration of follow up.

Table 8.7: Probability Density Functions for Parameters Relating to Information Value from Etidronate Trial

	Base value	Probability density function
Start up costs	$30 000	Normal (30 000, 3 000)
Annual costs of study management	$50 000	Normal (50 000, 5 000)
Costs per sample	$500	Normal (500, 50)
Finishing costs	$10 000	Normal (10 000, 1 000)
Analysis and policy costs	$10 000	Normal (10 000, 1 000)
Time required for start up	0.25	Triangular (0.15, 0.35, 0.35)
Time required for finishing	0.17	Triangular (0.15, 0.17, 0.19)
Time required for analysis and policy revision	0.17	Triangular (0.15, 0.17, 0.19)
Annual recruitment rates	365	Normal (365, 36.5)
Annual number of patients	74 848	Normal (74 848. 7 485)
Uptake of results	0.75	Beta (30, 10)
Lifespan of research	10	Triangular (7, 10, 13)

Note: Normal distributions characterized by mean and standard error. Triangular distributions characterized by minimum, mode, maximum. Beta distribution characterized by alpha and beta.

Figure 8.7 Frequency Count for Optimal Sample Size for Etidronate Clinical Assuming Data are Uncertain

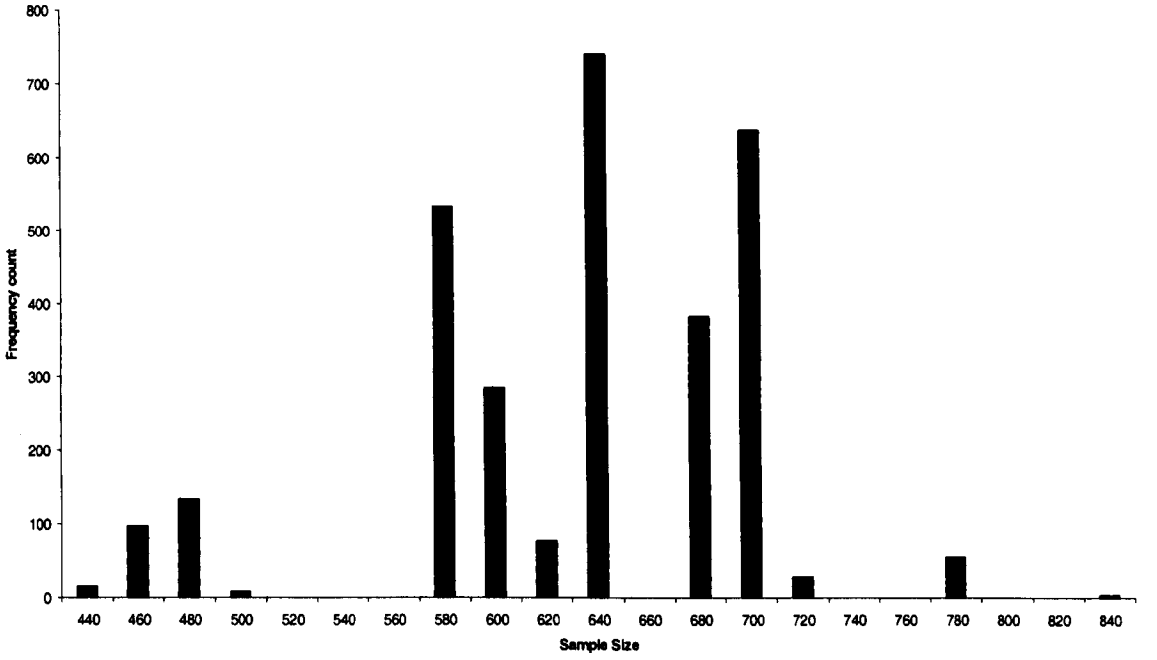
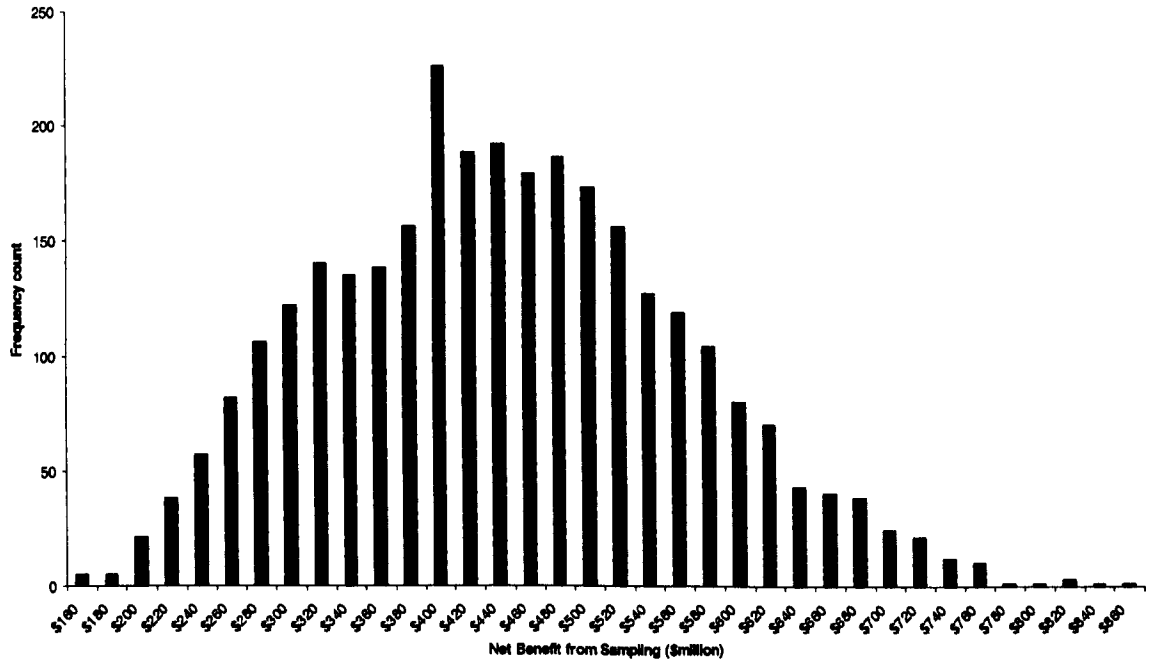


Figure 8.8 Frequency Count for Net Benefit of Sampling for Etidronate Clinical Assuming Data are Uncertain



for both the optimal sample size (632.0) and the net benefit from sampling ($43.3 million) are similar to the original values from the base analysis (640 and $43.6 million respectively) (Table 8.8) . However, the frequency distributions and the 95% certainty intervals from the Monte Carlo simulation highlight the uncertainty over these estimates.

8.4.2 Impact of changes to the design of clinical trials

The optimal sample size for the randomized controlled trials was based on particular assumptions relating to the design of the trial. Studies were assumed to be single site RCTs. In the preceding section, the impact of uncertainty around parameters relating to the original study design was assessed for the etidronate clinical trial.

In this section, the impact of conducting a multi centre clinical trial is assessed. A multi site study will allow quicker completion of study and thus increasing the duration for which a study may make an impact. However, a multi site study is likely to lead to increased costs relating to both start up and the annual costs associated with the coordination of data collection and investigator costs.

For the single site study, it was assumed that start up costs were $30 000 with annual costs of $50 000. To assess the impact of multi site studies, the following assumptions were made. First, each site would recruit patients at the same rate – i.e. a two site study would finish recruitment in half the time of the single site study. However, for each additional site participating in the study there would be a delay in the start of recruitment due to a necessary increase in start up time (assumed to be 1 month per additional site). Secondly, both start up costs and annual costs would

Table 8.8: Optimal Sample Size and Maximum Net Benefit of Sample Information for the Etidronate Clinical Trial

	Optimal Sample Size	Net Benefit from Sample Information
Base analysis	640	$43.6 million
Monte Carlo simulation	632.0	$43.3 million
	(460, 720)	(22.2 million, 67.8 million)

Note: Figures in parenthesis are 95% certainty intervals.

increase by $10 000 for each additional site involved reflecting increased management costs associated with multi site studies.

Thus, conducting multi site studies will involve a trade off between higher costs of research and possible greater information value due to its shorter duration.

Figure 8.9 demonstrates the effect of the number of sites participating in the study on both the optimal sample size and the net benefit gain from sampling. For both trials, adding additional sites increases the optimal sample size as more patients can be recruited quicker and the delay in use of the research is limited. However, for both trials the maximum net benefit of sample information will peak at a specific number of sites (for etidronate 7 and for alendronate 5). Adding additional sites would lead to a loss in net benefit as the increased costs outweigh the benefits form an increased sample.

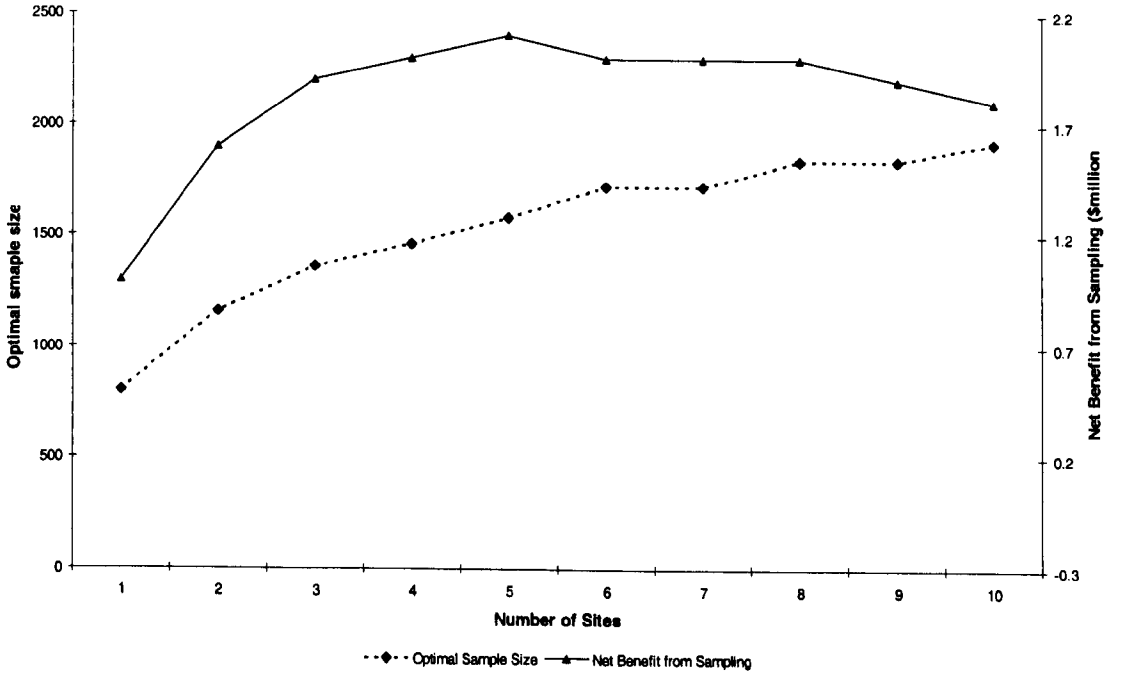
8.4.3 Randomised trial of no therapy, etidronate and alendronate

In Sections 8.3.5.1 and 8.3.5.2, the information value from conducting clinical trials comparing first, etidronate, and secondly, alendronate, to no therapy was assessed. Both trials generated significant net benefits which may be unsurprising given the relative uncertainty over which is the optimal treatment for the different sub-groups of patients.[78]

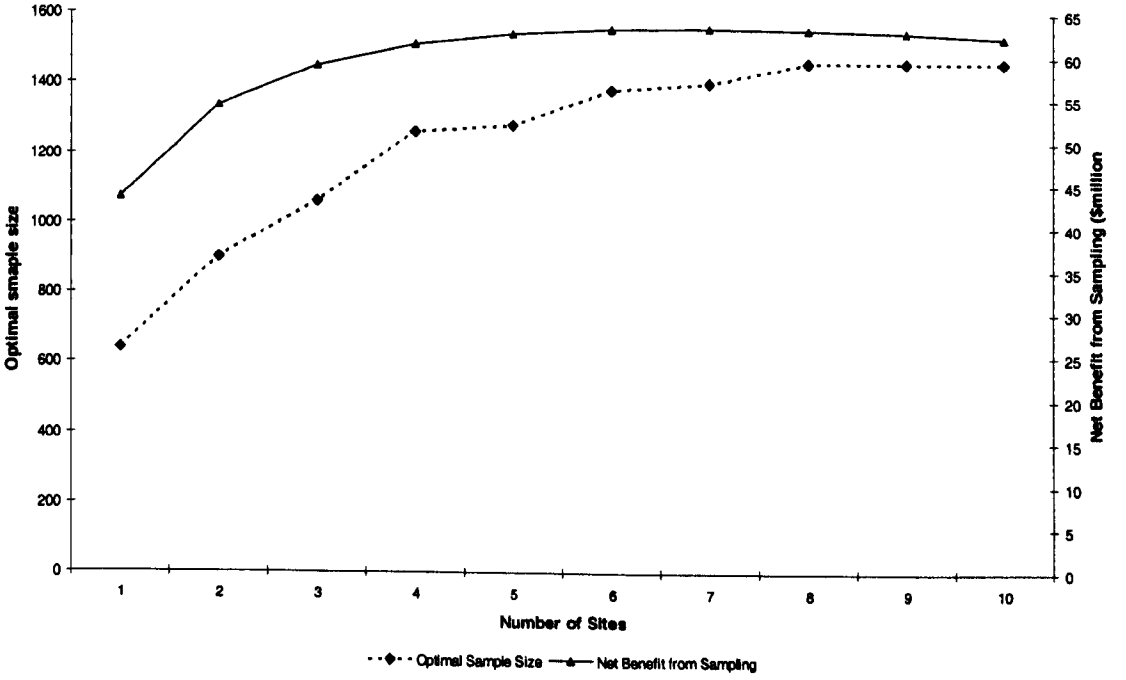
Both trials may be worthwhile. However, it is likely to be infeasible to conduct two simultaneous trials in the same patient population. As both trials include no therapy

[78] This is highlighted in Appendix B.

Figure 8.9 Effect of Number of Sites Participating in Clinical Trials on Optimal Sample Size and Net Benefit of Sample Information



a. Alendronate-no therapy clinical trial



b. Etidronate-no therapy clinical trial

as a comparator an alternate single three arm trial may be preferable. Thus, the methods used to assess the optimal sample size for the etidronate and alendronate trials were applied to the potential three arm clinical trial.

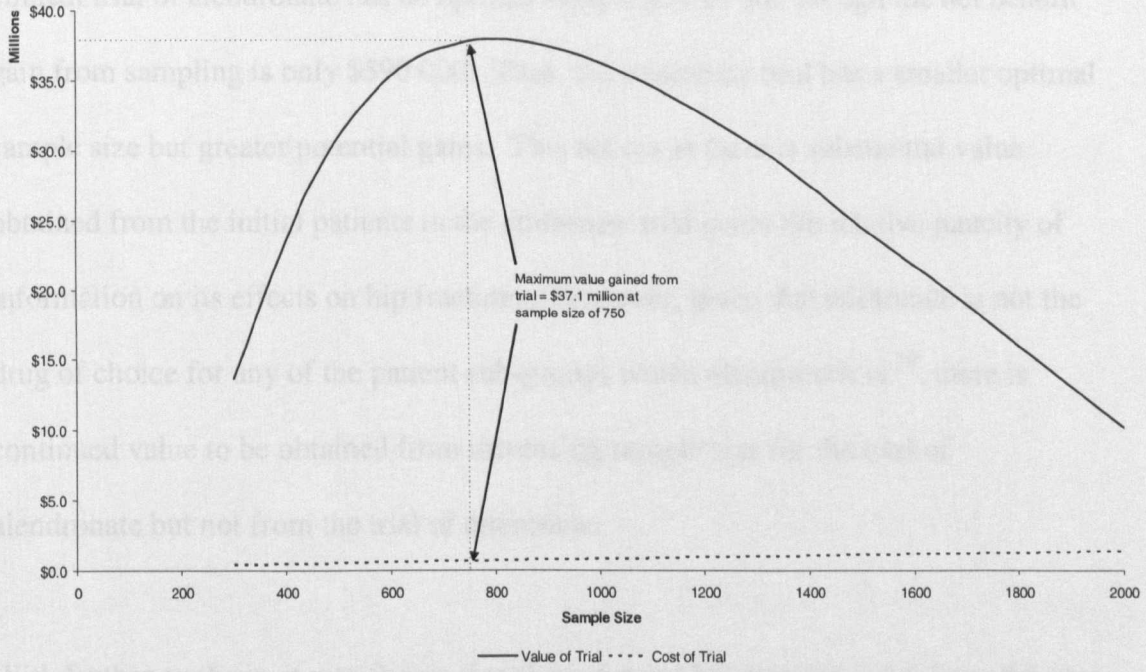
Figure 8.10 depicts the cost of sample information and EVSI by potential sample size for the three arm clinical trial. As with the etidronate-no therapy trial, EVSI is shown to rise significantly with information from the first few patients within a trial. For sample sizes greater than 800 EVSI is shown to fall as the value obtained from additional patients is less than the net benefit loss arising from the delay in obtaining information. The optimal sample size for the clinical trial is 750 patients with a net benefit of sample information of $37.1 million (Table 8.6).

8.5 CONCLUSIONS

In Section 8.2, the methods for assessing information value relating to treatment decisions for osteoporosis were described. In Section 8.3, the expected value of perfect partial information for individual parameters and parameter groups was estimated and potential studies which may have information value were identified. In Section 8.4 the implications on information value and optimal sample size of changes to some of the assumptions made for the analysis in Section 8.3 were assessed.

Six potential studies were identified. Two studies involved randomized controlled trials of treatment to assess their impact on hip fractures, one study involved a case control study to assess the relative increase in fractures with previous fracture and three further studies required simple cohort studies to refine individual parameters within the model.

Figure 8.10 EVSI and Cost of Sample Information for Trial including Etidronate, Alendronate and No Therapy by Sample Size



The greatest net benefit gain from sampling is from a randomized controlled trial examining the relative risk of hip fracture with etidronate. The maximum net benefit gain from this study is $30.3 million with a sample size of 640. The randomized clinical trial of alendronate has an optimal sample size of 800 though the net benefit gain from sampling is only $590 000. Thus, the etidronate trial has a smaller optimal sample size but greater potential gains. This occurs as there is substantial value obtained from the initial patients in the etidronate trial given the relative paucity of information on its effects on hip fractures. However, given that etidronate is not the drug of choice for any of the patient sub-groups whilst alendronate is [79], there is continued value to be obtained from increasing sample size for the trial of alendronate but not from the trial of etidronate.

With further analysis, it was shown that there was no information value from the case control study relating to the relative risk of fracture with previous fracture. The EVSI from this trial increases with sample size but is never greater than the estimated costs of the study. For the three potential cohort studies, there is net benefit to be gained from conducting these studies, though they are less than from the two potential clinical trials.

There is clear evidence that randomized controlled trials for etidronate and alendronate relating to their impact on preventing hip fractures may be justified given that they provide sufficient information value. However, the feasibility of conducting both trials needs to be questioned. In Section 8.4.3, an alternate approach is explored in terms of a three arm clinical trial involving no therapy, etidronate and

[79] As demonstrated in chapter 6.

alendronate. The EVPPI for the parameter group of the effect on hip fractures for alendronate and etidronate is by definition greater than the EVPPI for etidronate alone $337 versus $311. However, analysis shows that the three arm trial will have less information value than the trial comparing etidronate to no therapy. This is likely due to both the reduced proportion of patients in the trial receiving etidronate (and consequently the reduced level of information for the more uncertain parameter) and the delay in completion of the study. Thus, if one trial was to be conducted the optimal trial would be the comparison of etidronate to no therapy.

In Section 8.4.1, the impact of uncertainty around the parameters required for estimating optimal sample size and the net benefit of sampling was assessed. Uncertainty was shown to lead to a wide dispersion in estimates of these parameters. However, expected values were similar to estimates from the base analysis, implying that, in this instance, uncertainty in this area would have little impact on determining an efficient research plan.

In Section 8.4.2, analysis focused on identifying whether a multi site trial would be more beneficial than a single site trial. Involving more sites, requires the consideration of the trade off between the ability to recruit more patients quicker and increased research costs. Thus, analysis found that although optimal sample size will increase as the number of sites increase; there will be an optimum number of sites which will maximize the net benefit of sample information and this will vary by study.

There is a less convincing case for the conduct of the three potential cohort studies. The parameter for which there is the greatest evidence in favour of conducting a further study is the proportion of benefit maintained after curtailment of treatment. Analysis is based on the conduct of a hypothetical cohort study. However, this parameter is extremely difficult to measure. An alternate approach to measuring this parameter would be to conduct a long term follow up of patients within a clinical trial. This may be less costly but will result in longer delay before the data would become available. Data for the other two parameters relating to long term care can be obtained from specific cohort studies. However, for both parameters the net benefit gain is small and an alternate approach of obtaining further information from the proposed randomized controlled trials may be preferable.

In conclusion, in this chapter, value of information analysis has been conducted to identify an efficient research plan relating to the treatment of osteoporosis. Given the underlying uncertainty concerning the effectiveness of etidronate (the cheapest of the available bisphosphonates), it is not surprising that a clinical trial to obtain further information on the drug's effectiveness is the most desirable of the potential research studies. A clinical trial relating to alendronate is of less value but may be justified. However it may be infeasible to conduct both trials simultaneously and a three arm trial has been shown to be of less value than the simple two arm etidronate trial. Further information for three other parameters may be warranted. However, if a trial of etidronate was conducted it is possible that such information as well as information on many other parameters could be obtained for limited additional cost.

Chapter 9.

Conclusions

9.1 INTRODUCTION

The objectives of this thesis are: first, to develop a normative framework for fully handling both variability and uncertainty when making decisions using economic evaluation; and secondly, to apply this framework to an economic evaluation of treatments for osteoporosis. Decisions considered relate to, first, the immediate decision concerning which therapies should be funded for the treatment of which patients and, secondly, the decision concerning what is the optimal research plan to reduce uncertainty around the first decision.

In Chapter 2, the concepts of uncertainty and variability were defined and their previous handling in economic evaluation was discussed. In Chapters 3 and 4, the normative framework for handling uncertainty and variability was developed. In Chapter 5 and 6, the background to economic evaluation in osteoporosis was discussed and the model used in subsequent chapters was described. In Chapter 7, a stratified economic analysis of treatments for osteoporosis was conducted. In Chapter 8, a value of information analysis was conducted.

In this chapter, the conclusions to be drawn from this thesis are developed. In Section 9.2, the normative framework for handling uncertainty and variability are discussed. In Section 9.3, the feasibility and acceptability of the framework is discussed in relation to the evaluation of treatments of osteoporosis. Section 9.4

contains a discussion of the innovative features of this thesis are discussed. In Section 9.5, further work which could be conducted to enhance the progress made with the thesis is identified. Finally, Section 9.6 contains concluding remarks.

9.2 THE NORMATIVE FRAMEWORK

The primary objective of this thesis was to develop a normative framework for the consideration of uncertainty and variability when applying economic evaluations in health care decision making. The framework is argued to be intuitive in that unlike the previous practice of sensitivity analysis it is built on the actual definitions of both terms. In addition, the framework fully handles uncertainty and variability related to input parameters and outcomes.

Variability relates to the heterogeneity amongst the potential population affected by the choice of health care intervention. Given that it is not possible to reduce the degree of heterogeneity within the population the intuitive response is to split the population into groupings of more similar individuals. Thus, heterogeneity within each sub-group will be reduced.

Given the above, conducting economic analysis for different strata of the population is attractive in that it is known that the cost effectiveness of interventions may vary by patient characteristics. Thus, the practice of limiting interventions to those strata for which they are cost effective can realise substantive economic and health benefits for the population as a whole.

The technique of stratified cost effectiveness analysis is described in Chapter 3. The technique allows identification of the incremental net benefit to be gained from restricting access to interventions only to those individuals for which it is cost effective. The technique is feasible in that the outcomes generated are consistent with the format of recent decisions relating to the management of pharmaceuticals and devices in terms of limiting access. Furthermore, the technique permits consideration of both equity concerns and the impact of non-adherence.

Uncertainty relates to the lack of knowledge concerning a particular variable. As in economic evaluation we typically have a degree of uncertainty around parameter values, this propagates into uncertainty around the outcomes of interest; i.e. the relative cost effectiveness of alternate uses of scarce resources.

The framework is based upon the convincing arguments that in terms of public funding decisions the only outcomes of interest are the expected values. Thus, the thesis adopts what has been described as a Bayesian approach to decision making. It is important to note that the framework requires the conduct of probabilistic analysis to estimate expected values; recognising the non-linear relationships between input parameters and outcomes.

As uncertainty is caused by lack of perfect information then the collection of further information can reduce such uncertainty. Thus, the framework for handling uncertainty is twofold. First, decisions on funding should be based on expected values generated through probabilistic analysis. Secondly, that uncertainty around

the cost effectiveness of interventions should determine an efficient research plan relating to the particular treatment decision.

For the latter issue, the methods for assessing expected value of perfect information and the expected value of sample information detailed in Chapter 4 are argued to be appropriate for assessing the optimal sample size for potential research studies. The framework goes further than previous literature in that it allows the concurrent consideration of both uncertainty and variability. Thus, the optimal research plan relates to the need for further information to potentially revise the criteria for access to interventions.

The normative framework presented in this thesis allows decision makers to address the two fundamental questions arising from variability and uncertainty

1. Which patients should receive which therapies?
2. What further information should be collected to increase certainty concerning the answer to the previous question?

9.3 APPLICATION OF THE NORMATIVE FRAMEWORK

In the latter chapters of this thesis, the normative framework is applied to decisions relating to the appropriate interventions for osteoporotic women in Canada. Osteoporosis is a complex disease which subsequently requires the construction of a complex decision model incorporating various sub models relating to osteoporotic status, fractures, residence and mortality. The model is designed to determine optimal treatment choices for different patient cohorts and is fully probabilistic.

Analysis is time consuming; one Monte Carlo simulation of 3000 replications takes 4 ½ hours to run. For the analysis contained in Chapters 7 and 8, the osteoporosis model is subject to 643 separate Monte Carlo simulations of 3000 replications. Combined with the simulated research studies in section 8.3.5, the total number of replications required is 1.98 million. This, of course is the minimum number required assuming that no analytical errors will lead to repeat analysis!

Despite the associated computational complexity, the feasibility of applying the normative framework is demonstrated. Thus, the application of the normative framework in practice is argued to be both optimal and feasible though the computational complexity of the analysis is discussed further in Section 9.5.

Analysis determined the optimal limited use criteria for access to bisphosphonates. The stratified analysis has been shown to be pertinent to decision makers, in that, reflecting this analysis, the province of Ontario has revised the criteria for access to both alendronate and risedronate. Based on the results of the decision model, the revisions will lead to improved quality of life and survival for osteoporotic patients with an increase rather than a decrease in drug costs. However, whilst still sub optimal, it is clear that the revisions are based on the results of the analysis.

Analysis also determined an optimal research plan relating to treatment decisions for osteoporosis patients. However, it is less clear how the value of information analysis will influence decision makers. The major recommendation is to conduct a randomised clinical trial comparing etidronate with no therapy. Whilst, the optimal sample size of 640 is considerably less than that which could be justified based on

standard statistical inference, clinicians may not accept the need to conduct a clinical trial of the least effective treatment option particularly on ethical grounds. Thus, value of information analysis may be seen initially as informative rather than directive.

9.4 METHODOLOGICAL ADVANCES WITHIN THE THESIS

This thesis goes beyond recent work concerning the application of Bayesian techniques to economic evaluation, in that the thesis provides a complete framework for handling uncertainty and variability relating to input parameters within decision models. Previous recommendations tended to focus solely on uncertainty (Claxton et al. 2004).

The application of the framework to treatments for osteoporosis has many innovative features. It is the first study to fully assess uncertainty and variability within a Bayesian context. The analysis is the first study of treatments for osteoporosis which meets all previous recommendations for the conduct of studies in this area. The analysis includes one of the first full value of information analyses where a full range of parameters are considered and the value of a full range of potential study designs and sample sizes are assessed.

This thesis also includes several specific methodological developments. For handling variability in economic evaluation, the technique of stratified cost effectiveness analysis was developed which allows consideration of the net benefit gain from stratification. For handling uncertainty, two additional valid methods for assessing EVPPI are developed; the unit normal loss integral method and the quadrature

method. Furthermore, a previous method used to calculate EVPPI is shown to be invalid. Finally, unlike in previous analyses, the methods used for estimating the value of information considers both the delay in obtaining information due to the time required to conduct research and alternate sample sizes and study designs.

Recent changes to guidance relating to handling uncertainty and variability highlight that the recommendations contained within the normative framework are becoming recognized as appropriate. Recent guidance for the conduct of technology appraisals from the National Institute for Clinical Excellence (NICE 2004), contains stronger recommendations relating to the conduct of probabilistic sensitivity analysis. The guidance continues to recommend sub group analysis for handling variability with the focus on the need for clinical justification. Similarly, new guidelines currently being developed by CCOHTA are likely to include more advanced recommendations relating to uncertainty and variability[80]. However, these standards are far from routinely applied.

Nineteen HTA reports from the UK NHS HTA containing full economic evaluations based on decision analysis were published between January 2003 and May 2004 (the date of publication of the revised NICE Guidance)[81]. Of these 7 contained probabilistic sensitivity analysis (Bagnall et al. 2003, Chilcott et al. 2003b, Turner et al. 2003, Jones et al. 2004, Kalenthaler et al. 2004, Pandor et al. 2004, Barton et al. 2004). Only one study including estimation of EVPPI (Pandor et al. 2004). Only one

[80] Personal communication from Bruce Brady, CCOHTA.

[81] Bagnall et al. 2003, Barton et al. 2004, Calvert et al. 2003, Chilcott et al. 2003b, Clark et al. 2003, Davenport et al. 2003, Grimshaw et al. 2003, Hummel et al. 2003, Jones et al. 2004, Kalenthaler et al. 2004, Meads et al. 2003, Mowatt et al. 2003, Pandor et al. 2004, Roderick et al. 2003, Ross et al. 2004, Sharp et al. 2003, Song et al. 2003, Turner et al. 2003, Wardlaw et al. 2004

study contained a stratified analysis (Mowatt et al. 2003) with two further studies containing sub group analyses (Turner et al. 2003, Barton et al. 2004).

Thus, given previous economic evaluations, the revised recommendations from NICE are a major undertaking which will require considerable change to previous practice. However, with respect to the framework for handling uncertainty and variability developed in this thesis it could be argued that both the NICE and CCOHTA guidelines do not go far enough. Neither guideline requires the conduct of a full stratified analysis. Within the NICE guidelines consideration of sub groups continues to be based on clinical rather than economic criteria. Furthermore, the NICE guidelines focuses on presenting the results of probabilistic sensitivity analyses in the form of certainty intervals and cost effectiveness acceptability curves with only limited discussion relating to assessing the value of further information.

Thus, revisions to guidelines for economic evaluation are moving towards recommendations contained within this thesis. However, revisions are required before they meet the standards proposed in this thesis for the optimal handling of uncertainty and variability.

9.5 FURTHER WORK REQUIRED

Further work is still required before all issues relating to uncertainty and variability are covered.

The thesis does not consider uncertainty relating to the appropriate model structure. When developing decision models assumptions are required relating to many model

features such as cycle length, the number of health states and time horizon. These assumptions reflect an understanding of the underlying disease and treatment process for osteoporotic patients and are not specifically evidence based. The assumptions can influence the results of an analysis. The impact of assumptions can be assessed through simple sensitivity analysis but further advanced recommendations on handling such issues need to be made.

The thesis does not consider how to deal with any lack of data for stratified analysis. Within the economic evaluation of treatments for osteoporosis it is assumed that the relative risk reductions associated with therapies are consistent across all patient strata. This assumption is made due to the lack of data for each strata. The implications of making such an assumption and the validity of assuming the same expected value and variance across all strata are areas requiring further research. Implicit within this discussion is the need to consider the trade off between reducing heterogeneity through stratified analysis and the potential for increased decision uncertainty. Thus, as the potential patient population is split into more groups the information value from further studies may increase.

The form of analysis recommended by the normative framework is complex. Although, the feasibility of using this framework within a complex decision model as required for osteoporosis was demonstrated in Chapters 7 and 8, the complexity of the analysis and the time required may limit its applicability in all settings. Thus, techniques for simplifying the conduct of repeated Monte Carlo simulations may be beneficial and the development of specific software would be welcome. Similarly,

guidelines on determining the optimal number of replications with a Monte Carlo simulation would be worthwhile.

Finally, the thesis highlights the impact of uncertainty around assumptions relating to information value. When estimating the value of further information, data are required relating to the likely characteristics of research, the likely uptake of research, the useful life span of research information and the potential patient population affected by any decisions relating to reimbursement.

- For design features of research, the development of standardised methods for estimating relevant parameters would be worthwhile. Also, the impact of assuming that such variables are uncertain rather than known should be addressed.
- Within this thesis, physician adherence to revised reimbursement decisions is assumed to be independent of the form of information which leads to the revision. It may be expected that physicians would be more likely to adhere to a change in reimbursement status which involved an expansion of the indication of a treatment rather than a contraction. Similarly, physicians may be more likely to adhere to revisions based on new clinical trial evidence rather than updated cost data.
- The life span of research information should be assumed to vary by form of input parameter. For example, clinical trial evidence is useful until the treatments consider within the trial are no longer considered suitable. Utility values for common health states could be considered as having a long lifespan given that they should be appropriate unless there are fundamental changes in society's attitudes to the health state. Finally, cost parameters

may have a limited life span unless it can be assumed there will not be any fundamental changes in how patients are managed over time

- Finally, in value of information analysis, it is common to consider only the value of information to the decision makers for the specific location to which the study relates. However, it is clear that research relating to clinical efficacy and possibly utilities will be of use to decision makers internationally. Thus, whether such external value should be considered will be dependent on whether decision makers consider such value of importance may depend on whether decision makers are involved in formal collaborations with respect to research funding.

9.6 CLOSING REMARKS

The primary objective of this thesis was to develop a normative framework for handling uncertainty and variability concerning input parameters within economic evaluation based on decision analysis. It is argued that the framework developed in Chapters 3 and 4 provides the optimal methods for handling these concepts when making decisions regarding treatment reimbursement. First, the framework requires the stratification of patients into cohorts which are more homogenous with respect to costs and outcomes. Secondly, the framework involves the determination of an optimal research plan to increase the certainty relating to the above decisions.

As outlined in Section 9.4, the thesis contains several methodological advances relating to the analysis of uncertainty and variability. These relate to the methods for conducting stratified cost effectiveness analysis and methods for estimating EVPPI. The framework allows the full consideration of uncertainty and variability with respect to input parameters. There are other aspects of uncertainty which still require

similar consideration. In the previous section, further research is identified which would enhance the framework further by allowing consideration of other aspects relating to uncertainty and variability.

The conduct of an economic evaluation of treatments for osteoporosis using the normative framework detailed in Chapters 7 and 8 highlights the feasibility of conducting such studies. Furthermore, recent changes to reimbursement decisions within Ontario demonstrate the acceptability of this framework.

The thesis provides a blueprint for future studies in terms of handling uncertainty and variability through the adoption of the normative framework. The thesis contains the first example of an economic analysis fully considering uncertainty and variability. The recommendations from this analysis provide both the optimal treatment choices for specific patient cohorts as well as an optimal research plan. It is hoped that future studies will adhere to the standards developed in this thesis.

APPENDIX A: COST EFFECTIVENESS ACCEPTABILITY CURVES WITH MULTIPLE TREATMENT OPTIONS

A.1 Introduction

Cost effectiveness acceptability curves (CEAC) present for different values of λ, the probability that each treatment is the most cost effective given the available evidence (see Section 3.4.4).

In a two treatment model, a CEAC will report the percentage of replications where one treatment is optimal, the complement of this percentage is the percentage of replications where the other treatment is optimal. However, as demonstrated by the entacapone case study in Chapter 3, at specific values of λ, the optimal treatment as determined by the expected value of net benefit need not have the higher probability of being cost effective. Thus, the use of CEACs seems limited to a simple method of graphically representing the degree of uncertainty around the optimal treatment and not as a means of facilitating a decision.

In a multiple treatment model, the CEAC curve for each treatment can be represented on one graph (the sum of the height of all curves at each value of λ will be 1). In this instance, the y axis represents the proportion of replications where each treatment is associated the maximum NB. In this appendix, issues concerning the interpretation of CEACs with multiple treatment options are identified by focusing on one particular strata from the analysis of osteoporosis treatments : an 80 year old woman with previous fracture.

A.2 Methods

Multiple estimates of costs and QALYs for each osteoporotic treatment option were derived for the strata relating to an 80 year old woman with previous fracture. The expected values were used to determine the cost effectiveness frontier for the strata (i.e. which therapies are optimal at which values of λ).

A CEAC was drawn based on the multiple estimates of costs and QALYs. In addition, a CEAC was derived assuming only alendronate and etidronate were the only available therapeutic options.

A.3 Results

Table A.1 presents the expected values for each therapeutic option. Etidronate is dominated by all other treatment options whilst risedronate is dominated by alendronate. Thus, the optimal treatment choice is either no therapy or alendronate dependent on the value of λ: for $\lambda < \$11\,600$ no therapy is optimal, for all other others alendronate.

There is high degree of uncertainty over the true value of costs and QALYs especially for etidronate. This is depicted in Figure A.1 which is a scatterplot of the estimates of incremental costs and QALYs for each bisphosphonate compared to no therapy.

This high degree of uncertainty, over the costs and effects associated with etidronate leads to the shape of the cost effectiveness acceptability curve in Figure A.2. Despite etidronate being dominated by all other treatment options, it is the treatment option

with the highest probability of being cost effective for values of λ both < $20 000 and > $100; and the second highest for all other values.

Figure A.3 depicts the probability of etidronate being more cost effective than alendronate. For all values of λ, alendronate has the highest probability of being cost effective.

A.4 Conclusions

CEACs present the probability that a specific treatment is the most cost effective. CEACs ignore the dispersion of values for different treatment alternatives; the skewness of the distribution of net benefit may mean that therapies that are optimal based on expected values but will not be the majority decision. Thus, CEACs can not be used to facilitate decision making concerning the optimal treatment choice.

Treatments for which there are more information available may have a lower probability of being preferred than treatments for which there is less information. This is demonstrated by the example above, in that etidronate which is dominated by all therapies but yet has the highest probability of being cost effective for certain values of λ. A more useful graphical depiction would be to compare individual therapies to each other rather than include all therapies within the same graph.

If CEACs were used as a means of choosing the optimal treatment choice, then it may not be beneficial for companies to collect more data on a product – for example, with the data below it would appear unwise for the makers of etidronate to collect more data as this would likely lead to less uncertainty around its cost effectiveness

and will make it less attractive based on a CEAC. Given the results from this study, the role of CEACs in aiding decisions must then be questioned.

Table A.1 Expected Values for Costs and QALYs for an 80 Year-old Woman with Previous Fracture

	Etidronate	No therapy	Risedronate	Alendronate
Lifetime cost	$9 600	$8 700	$9 400	$8 900
QALYs	4.272	4.283	4.293	4.305
Incremental cost per QALY				
vs. etidronate		Dominant	Dominant	Dominant
vs. no therapy			$67 600	$11 600
vs. risedronate				Dominant

Figure A.1 Scatterplot of Incremental Costs and QALYs

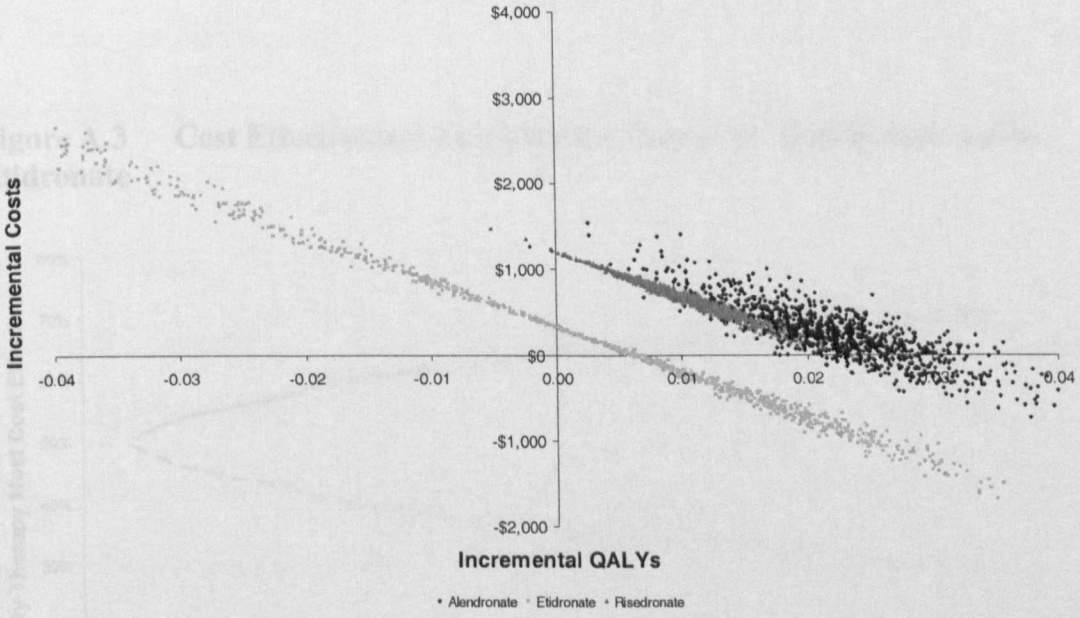


Figure A.2 Cost Effectiveness Acceptability Curve for Multiple Treatment Options

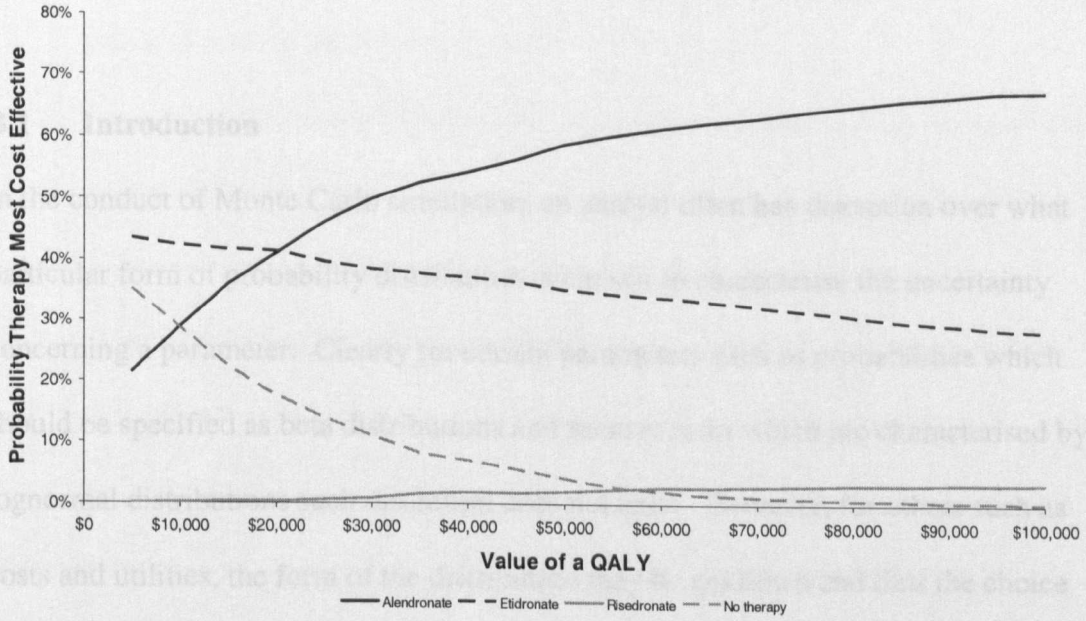
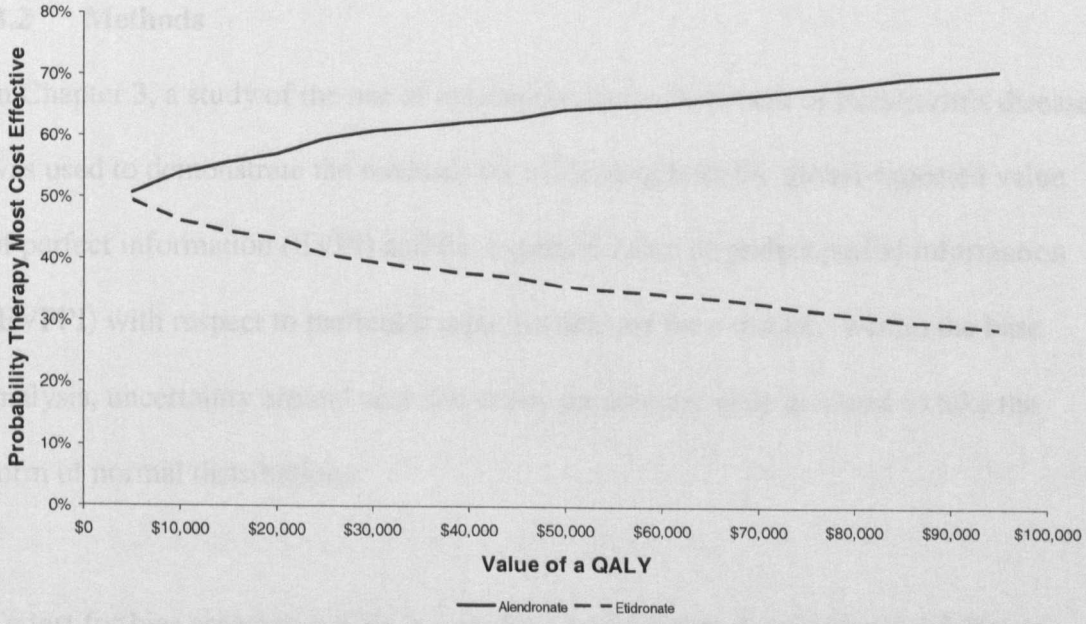


Figure A.3 Cost Effectiveness Acceptability Curve for Alendronate versus Etidronate



APPENDIX B: EFFECT OF CHOICE OF PROBABILITY DISTRIBUTION ON UNCERTAINTY WITHIN ECONOMIC ANALYSIS

B.1 Introduction

In the conduct of Monte Carlo simulation, an analyst often has discretion over what particular form of probability distribution is chosen to characterise the uncertainty concerning a parameter. Clearly for certain parameters such as probabilities which should be specified as beta distributions and relative risks which are characterised by lognormal distributions such discretion does not exist. However, for others such as costs and utilities, the form of the distribution may be unknown and thus the choice of distribution may bias results. Thus, the objective of this appendix is to identify the potential degree of bias inherent in the subjective choice of probability distributions

B.2 Methods

In Chapter 3, a study of the use of entacapone in the treatment of Parkinson's disease was used to demonstrate the methods for estimating both the global expected value of perfect information (EVPI) and the expected value of perfect partial information (EVPPI) with respect to particular input parameters for a model. Within the base analysis, uncertainty around cost and utility parameters were assumed to take the form of normal distributions.

To test for bias associated with this choice of probability distribution the following was conducted.

1. The mean and standard error for both costs and disutilities from the original model were identified.

2. For both costs and utilities uncertainty was expressed in terms of normal (original), gamma and lognormal distributions with the same mean and standard error as the base analysis
3. For each of the nine (3x3) possible combination of distributions the mean and 95% CI for incremental net benefit was identified
4. CEACs were plotted for each of the nine combinations
5. Global EVPI and EVPPI for parameter groups of costs and disutilities were estimated for each different combination

B.3 Results

Table B.1 presents estimates of the incremental net benefit for entacapone based on the different combinations of distributions. Expected values are identical though the width and skewness of the 95% certainty interval vary; although minimally. The maximum width of the interval is $9 089 associated with normal distributions for costs and gamma distributions for disutilities. The minimum width of $9 066 is associated with lognormal distributions for costs and gamma distributions for disutilities. The shape of the CEAC is consistent across all combinations of probability distributions (Figure B.1).

Estimates of global EVPI are consistent across combinations ranging from $147 and $152 per patient. However, estimates of EVPPI are less consistent (Table B.2). For costs, EVPPI varies from $3.30 for lognormal distributions to $7.05 for normal distributions. For disutilities, EVPPI varies from $24.20 for lognormal distributions to $26.18 for normal distributions.

B.4 Conclusions

Given that costs and utilities are weights to be applied to estimates of time duration and incidence, the results are as expected in that the choice of distribution has no effect on the expected value of net benefit. The choice of distribution has a modest effect on the 95% CI of NB and on global EVPI. However, the choice has a much more noticeable effect on estimates of EVPPI which in turn will affect EVSI and optimal sample size. However, the direction of the bias is not consistent across the two parameter groups.

Thus, in the conduct of value of information analysis, analysts must be careful in their choice of parameter distribution.

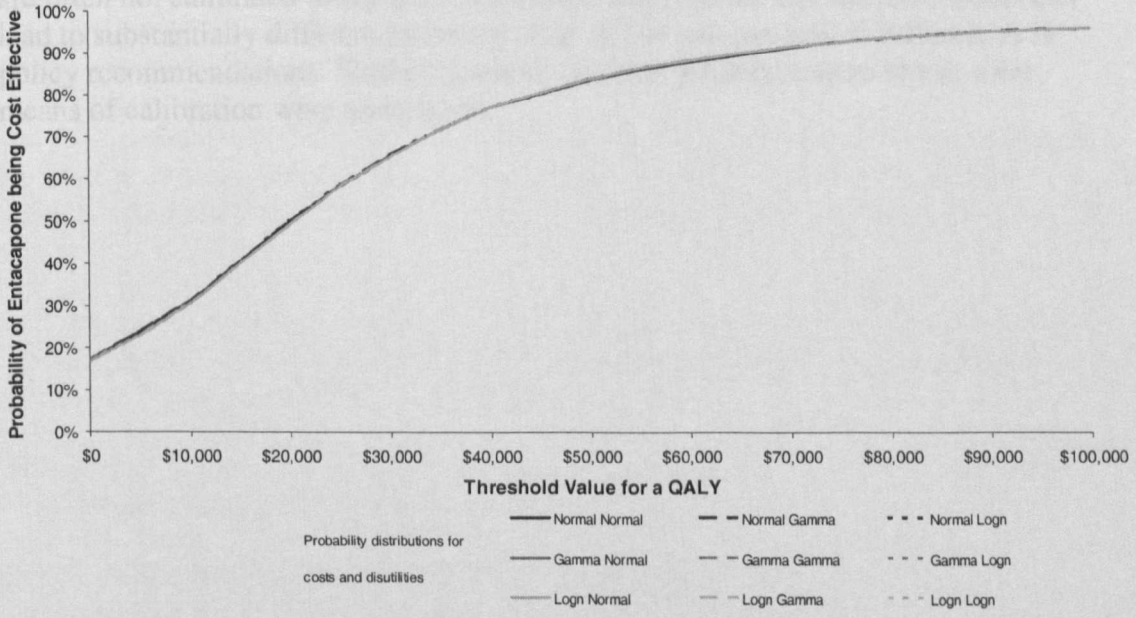
Table B.1 Incremental Net Benefit and EVPI by Choice of Probability Distribution

Cost Distribution	Disutilities distribution	Net Benefit	95% CI	EVPI
Normal	Normal	2252	-1715, 7368	151
Gamma	Normal	2252	-1687, 7385	149
Lognormal	Normal	2252	-1675, 7394	147
Normal	Gamma	2252	-1718, 7371	152
Normal	Lognormal	2253	-1717, 7362	152
Gamma	Gamma	2253	-1694, 7380	149
Gamma	Lognormal	2253	-1699, 7378	150
Lognormal	Gamma	2252	-1680, 7386	148
Lognormal	Lognormal	2253	-1687, 7385	148

Table B.2 EVPPI for Parameter Groups by Choice of Probability Distribution

Parameter Set	Distribution	EVPPI
Costs	Normal	7.05
Costs	Gamma	3.95
Costs	Lognormal	3.03
Disutilities	Normal	24.20
Disutilities	Gamma	25.28
Disutilities	Lognormal	26.18

Figure B.1: Cost Effectiveness Acceptability Curves by Choice of Probability Distributions



APPENDIX C: ABSTRACT ACCEPTED FOR PRESENTATION AT THE 2004 MEETING OF THE SOCIETY FOR MEDICAL DECISION MAKING

The Impact of Failure to Calibrate on Results in Economic Evaluation

Douglas Coyle, Ottawa Health Research Institute, Ottawa, Canada

Purpose: To demonstrate how failure to calibrate economic models can lead to biased estimates of the cost effectiveness of health interventions

Methods: The Canadian Economic Model of Osteoporosis has been used in several previous evaluations of osteoporosis treatments. It is fully calibrated in that the model replicates population data for the risk of fracture and mortality. Many previous models of osteoporosis are not calibrated: e.g. the population risk of fractures are weighted by the relative risks of fracture with osteoporosis or previous fracture history. This will lead to an overestimation of the risk of fracture for such groups. Analysis assessed the cost effectiveness of alendronate compared to no therapy for a 75 year old women with previous fracture history with and without calibration for fracture risk and mortality. Analysis identified the optimal age at which treatment with alendronate becomes cost effective.

Results: For a 75 year women with previous fracture, the annual probability of fracture without therapy is 1.31% with calibration and 1.58% without calibration. The incremental cost per QALY gained (ICUR) of alendronate is $35 600 when calibrating the model with respect to fracture risks and mortality and $19 400 without. Assuming a QALY was worth $50000, it would be cost effective to treat with alendronate women with previous fracture who were aged 75 and over when the model was calibrated. Without calibration, alendronate would be cost effective for women aged 73 and over.

Conclusions: Although recommended, decision models used for economic analysis are often not calibrated to replicate population data. Failure to calibrate models can lead to substantially different estimates of an ICUR and can lead to differences in policy recommendations. Studies based on decision models need to report what means of calibration were undertaken.

REFERENCES

Ades AE, Lu G, Claxton K. (2004) Expected value of sample information calculations in medical decision modeling. Med Decis Making, vol. 24, pp. 207-227.

Altman DG, Machin D, Bryant TN, Gardner MJ. (2000) Statistics with Confidence, London, BMJ Books.

Aminoff MJ. (1998) Parkinson's disease and other extrapyramidal disorders. In: Fauci AS, Braunwald E, Isselbacher KJ, et al. (ed.). Harrison's Principles of Internal Medicine 14 ed, New York, NY: McGraw-Hill - Health Professions Division.

Ankjaer-Jensen A, Johnell O. (1996) Prevention of osteoporosis: cost-effectiveness of different pharmaceutical treatments. Osteoporos Int, vol. 6, pp. 265-75.

Arrow KJ, Lind RC. (1970) Uncertainty and the evaluation of public investment decisions. American Economic Review, vol. 60, pp. 364-78.

Aursnes I, Storvik G. Gåsemyr J, Natvig B (2000). A Bayesian analysis of bisphosphonate effects and cost-effectiveness in postmenopausal osteoporosis. Pharmacoepidemiology and Drug Safety, vol. 9, pp. 501-509.

Bagnall AM, Jones L, Richardson G, Duffy S, Riemsma R. (2003) Effectiveness and cost-effectiveness of acute hospital-based spinal cord injuries services: systematic review. Health Technol Assess, vol. 7, no.19.

Baker RD. (2002) Sensitivity analysis for healthcare models fitted to data by statistical methods. Health Care Manage Sci vol. 5, pp. 275-281.

Baltussen RM, Hutubessy RC, Evans DB, Murray CJ. (2002) Uncertainty in cost-effectiveness analysis. Probabilistic uncertainty analysis and stochastic league tables. Int J Technol Assess Health Care, vol. 18, pp. 112-9.

Bartell SM, Breck JE, Gardner RH, Brenkert AL. (1986) Individual parameter perturbation and error analysis of fish bioenergetics models. Can J Fish Aquat Sci, vol.43, pp. 160-168.

Barton P, Jobanputra P, Wilson J, Bryan S, Burls A. (2004) The use of modelling to evaluate new drugs for patients with a chronic condition: the case of antibodies against tumour necrosis factor in rheumatoid arthritis. Health Technol Assess, vol. 8, no.11.

Belgian Society of Pharmacoepidemiology. (1995) A proposal for methodological guidelines for economic evaluation of pharmaceuticals-Belgium. Gent, Belgian Society of Pharmacoepidemiology (BESPE).

Bendich A, Leader S, Muhuri P. (1999) Supplemental calcium for the prevention of hip fracture: potential health-economic benefits. Clin Ther, vol. 21, pp. 1058-72.

Best L, Milne R. (1998) Bisphosphonates in the management of osteoporosis. Southampton: Wessex Institute for Health Research & Development, Development & Evaluation Report No. 79.

Birnbaum, ZW. (1969) On the importance of different components in a multicomponent system, in: PR Krishnaiah, ed., Multivariate Analysis-II, Academic Press, New York.

Black D, Cummings S, Karpf D, Cauley J, Thompson D, Nevitt M et al. (1996) Randomised trial of effect of alendronate on risk of fracture in women with existing vertebral fractures. Fracture Intervention Trial Research Group. Lancet, vol. 348, pp. 1535-41.

Black WC. (1990) The CE plane: a graphic representation of cost-effectiveness. Med Decis Making, vol. 10, pp. 212-4.

Bleichrodt H. (1997) Health utility indices and equity considerations. Journal of Health Econ vol. 16, pp. 65-91.

Brazier JE, Green C, Kanis JA. (2002) A systematic review of health state utility values for osteoporosis-related conditions. Osteoporos Int, vol. 13, pp. 768-776

Brennan AB, Chilcott JB, Kharroubi S, O'Hagan A. (2002a) Calculating expected value of perfect information:- resolution of conflicting methods via a two level Monte Carlo approach. 24th Annual Meeting of SMDM, Washington.

Brennan AB, Chilcott JB, Kharroubi S, O'Hagan A. (2002b) A two level Monte Carlo approach to calculation expected value of sample information: how to value a research design. 24th Annual Meeting of SMDM, Washington.

Briggs A, Fenn P. (1998) Confidence intervals or surfaces? Uncertainty on the cost-effectiveness plane. Health Econ, vol. 7, pp. 723-40.

Briggs A, Goeree R, Blackhouse G, O'Brien B. (2002) Probabilistic analysis of cost-effectiveness models: choosing between treatment strategies for gastro-esophogeal reflux disease. Med Decis Making, vol. 22, pp. 290-308.

Briggs A, Sculpher M, Buxton M. (1994) Uncertainty in the economic evaluation of health care technologies - the role of sensitivity analysis. Health Econ vol. 3, pp. 95-104.

Briggs AH, Wonderling DE, Mooney CZ. (1997) Pulling cost-effectiveness analysis up by its bootstraps: a non-parametric approach to confidence interval estimation. Health Econ, vol. 6, pp. 327-40.

Briggs AH. (2000) Handling uncertainty in cost-effectiveness models. Pharmacoeconomics, vol. 17, pp. 479-500.

Brooks DJ, Forsyth D, Playfer JR, Williams AC. (2000) The role of entacapone in the management of Parkinson's disease. Hospital Medicine (London), vol. 61, pp. 267-71.

Calvert N, Hind D, McWilliams RG, Thomas SM, Beverley C, Davidson A. (2003) The effectiveness and cost-effectiveness of ultrasound locating devices for central venous access: a systematic review and economic evaluation. Health Technol Assess, vol. 7, no.12.

Canadian Coordinating Office for Health Technology Assessment. (1997) Guidelines for economic evaluation of pharmaceuticals: Canada. Ottawa, CCOHTA.

Canadian Pharmacists Association. (2001) Compendium of Pharmaceuticals and Specialties. 36th ed. Ottawa, Canadian Pharmacists Association.

Chaudhary MA, Stearns SC. (1996) Estimating confidence intervals for cost-effectiveness ratios: an example from a randomized trial. Stat Med, vol. 15, pp. 1447-58.

Cheok MC, Parry GW, Sherry RR. (1998) Use of importance measures in risk informed applications, Reliab Eng and Sys Safety, vol. 60, pp. 213-226.

Chesnut CH, McClung MR, Ensrud KE, Bell NH, Genant HK, Harris ST et al. (1995) Alendronate treatment of the postmenopausal osteoporotic woman: effect of multiple dosages on bone mass and bone remodeling. Am J Med, vol. 99, pp. 144-52.

Chilcott J, Brennan A, Booth A, Karnon J, Tappenden P. (2003a) The role of modelling in prioritising and planning clinical trials. Health Technol Assess, vol. 7, pp. 1-125.

Chilcott J, Lloyd Jones M, Wight J, Forman K, Wray J, Beverley C, Tappenden P. (2003b) A review of the clinical effectiveness and cost-effectiveness of routine anti-D prophylaxis for pregnant women who are rhesus-negative. Health Technol Assess, vol. 7, no. 4.

Choi, J, Harvey JW, Conklin MH. (1999) Use of multi-parametric sensitivity analysis to determine relative importance of factors influencing natural attenuation of mining contaminants. In Morganwalp DW, Buxton HT (eds) U.S. Geological Survey Toxic Substances Hydrology Program--Proceedings of the Technical Meeting, Charleston, South Carolina, Vol 1 –Contamination from Hardrock Mining.

Clark W, Raftery J, Song F, Barton P, Cummins C, Fry-Smith A, Burls A. (2003) Systematic review and economic evaluation of the effectiveness of infliximab for the treatment of Crohn's disease. Health Technol Assess, vol. 7, no. 3.

Claxton K, Fenwick E, Palmer S, Sculpher M, Abrams K, Sutton A. (2004) Building a Reference Case for Bayesian Applications to Health Economics and Outcomes Research. Accessed August 25th 2004 from http://www.bayesian-initiative.com/data/Claxton%20Final%20Report%20Introduction.pdf

Claxton K, Neumann PJ, Araki S, Weinstein MC. (2001) Bayesian value of information analysis. An application to a policy model of Alzheimer's disease. Int J Technol Assess Health Care, vol. 17, pp. 38-55.

Claxton K, Posnett J. (1996) An economic approach to clinical trial design and research priority-setting. Health Econ, vol. 5, pp. 513-24.

Clemens K, Townsend R, Luscombe F, Mauskopf J, Osterhaus J, Bobula J. (1995) Methodological and conduct principles for pharmacoeconomic research. Pharmacoeconomics vol. 8, pp. 169-74.

Clemmesen B, Ravn P, Zegels B, Taquet AN, Christiansen C, Reginster JY. (1997) A 2-year phase II study with 1-year of follow-up of risedronate (NE-58095) in postmenopausal osteoporosis. Osteoporos Int, vol. 7, pp. 488-95.

Committee for Proprietary Medicinal Products (CPMP). (2001) Note for guidance on postmenopausal osteoporosis in women. European Agency for the Evaluation of Medicinal Products, London.

Commonwealth Department of Human Services and Health. (1995) Guidelines for the pharmaceutical industry on preparation of submissions to the Pharmaceutical Benefits Advisory Committee including major submissions involving economic analyses. Canberra: Australian Government Publishing Service.

Consensus Development Conference (1993): Diagnosis, prophylaxis and treatment of osteoporosis. Am J Med, vol. 94, pp. 646-50.

Cooper C, Atkinson EJ, Jacobsen SJ, O'Fallon WM, Melton LJ. (1993) Population-based study of survival after osteoporotic fracture. Am J Epidemiol, vol. 137 pp. 1001-1005.

Cooper C. (1993) Population-based study of survival after osteoporotic fracture. Am J Epidemiol, vol. 137, pp. 1001-5.

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

Coyle D, Barbeau M, Guttman M, Baladi J-F (2003b). The economic evaluation of pharmacotherapies for Parkinson's disease treatment. Parkinsonism Relat Disord, vol. 9, pp. 301-7.

Coyle D, Buxton MJ, O'Brien B. (2003a) Stratified cost-effectiveness analysis: a framework for establishing efficient limited use criteria, Health Econ, vol. 12, pp. 421-7.

Coyle D, Buxton MJ, O'Brien BJ. (2003c) Measures of importance for economic analysis based on decision modeling. J Clin Epidemiol, vol. 56, pp.989-97.

Coyle D, Cranney A, Lee KM, Welch V, Tugwell P. (2001a) Cost effectiveness of nasal calcitonin in postmenopausal women: use of Cochrane Collaboration methods for meta-analysis within economic evaluation. Pharmacoeconomics, vol. 19, pp. 565-75.

Coyle D, Cranney A, Tugwell P. (2003d) Economic evaluation of norethisterone acetate/ethinylestradiol (FemHRT) for women with menopausal symptoms. Pharmacoeconomics, vol. 21, pp. 661-9.

Coyle D, Lee KM. (2002) Evidence based economic evaluation: how the use of different data sources can impact results. In: Donaldson C, Mugford M, Vale L. (eds.) Evidence Based Health Economics, London, BMJ Books.

Coyle D, Maunsell E, Wells GA, Graham I, Lee KM, Meunier N, Peterson JE. (1999) Differences in preference values between cultural groups: women's preferences over lifetime risks associated with postmenopausal osteoporosis therapy. 15th Annual Meeting of the International Society of Technology Assessment in Health Care. Abstract no. 81.

Coyle D, Tosteson ANA. (2003) Towards a reference case for economic evaluations of osteoporosis treatments. J Rheumato, vol. 68, pp. 31-36.

Coyle D. (2003a) Determining the optimal combinations of mutually exclusive interventions: a response to Hutubessy and colleagues. Health Econ, vol 12, pp 159-62.

Coyle D. (2003b) WHO's better not best: appropriate probabilistic uncertainty analysis. Int J Technol Assess Health Care, Vol 19, pp 540-5.

Cranney A, Coyle D, Pham BA et al. (2001a) The psychometric properties of patient preferences in osteoporosis. J Rheum, vol. 28, pp. 132-7.

Cranney A, Coyle D, Welch V, Lee KM, Tugwell P. (1999) A review of economic evaluation in osteoporosis. Arthritis Care Res, vol. 12, pp. 425-34.

Cranney A, Guyatt G, Griffith L. (2002a) Summary of meta analyses of therapies for postmenopausal osteoporosis. Endocr Rev, vol. 23, pp. 570-578.

Cranney A, Guyatt G, Krolicki N, et al. (2001b) A meta-analysis of etidronate for the treatment of postmenopausal osteoporosis. Osteoporos Int, vol. 12, pp.140-51.

Cranney A, Tugwell P, Adachi J. (2002b) Meta analysis of risedronate for the treatment of postmenopausal women. Endoc Rev, vol. 23, pp. 517-523.

Cranney A, Wells G, Willan A. (2002c) Meta analysis of alendronate for the treatment of postmenopausal women. Endoc Rev, vol. 23, pp. 508-516.

Cree M, Soskolne CL, Belseck E et al. (2000) Mortality and institutionalization following hip fracture. J Am Geriatr Soc, vol. 48, pp. 283-288.

Crystal Ball (2000) version 5.1. Denver, Colorado; Decisioneering.

Cullen AC. (1995) The sensitivity of probabilistic risk assessment results to alternative model structures: a case study of municipal waste incineration. Journal of Air and Waste Management Association, vol. 45, pp. 538-546.

Cummings SR, Black DM, Thompson DE et al. (1998) Effect of alendronate on risk of fracture in women with low bone density but without vertebral fractures - Results from the fracture intervention trial. JAMA, vol. 280, pp. 2077-82.

Cummings SR, Karpf DB, Harris F, Genant HK, Ensrud K, LaCroix AZ, Black DM. (2002) Improvement in spine bone density and reduction in risk of vertebral fractures during treatment with antiresorptive drugs. Am J Med, vol. 112, pp. 281-289.

Dakins ME, Toll JE, Small MJ. (1994) Risk-based environmental remediation: decision framework and role of uncertainty. Environ Toxicol Chem, vol. 13, pp. 1907-1915.

Davenport C, Elley K, Salas C, Taylor-Weetman CL, Fry-Smith A, Bryan S, Taylor R. (2003) The clinical effectiveness and cost-effectiveness of routine dental checks: a systematic review and economic evaluation. Health Technol Assess, vol. 7, no. 7.

Department of Health (2001). National Service Framework for Older People 2001, London, Department of Health.

Dere W, Avouac B, Boers M, et al. (1998) Recommendations for health economics analysis to be performed with a drug to be registered in prevention or treatment of osteoporosis. Calcif Tissue Int, vol. 63, pp. 93-7.

Detsky AS. (1993) Guidelines for economic analysis of pharmaceutical products : A draft document for Ontario and Canada. Pharmacoeconomics, vol. 3, pp. 354-361.

Dolan P. (1998) The measurement of individual utility and social welfare. J Health Econ, vol. 17, pp.39-52.

Doubilet P, Begg CB, Weinstein MC, Braun P, McNeil BJ. (1985) Probabilistic sensitivity analysis using Monte Carlo simulation: a practical approach. Med Decis Making, vol. 5, pp. 157-177.

Drummond MF, Jefferson TO. (1996) Guidelines for authors and peer-reviewers of economic submissions to the British Medical Journal. BMJ, vol. 313, pp. 275-383.

Drummond MF, O'Brien BJ, Stoddart GI, Torrance GW. (1997) Methods for the economic evaluation of health care programmes. New York, Oxford University Press.

Eisenberg NA, Sagar B. (2000) Importance measures for nuclear waste repositories. Reliab Eng and Sys Safety, vol. 70, pp. 217-239.

Felli JC, Hazen GB. (1998) Sensitivity analysis and the expected value of perfect information. Med Decis Making, vol. 18, pp. 95-109.

Felli JC, Hazen GB. (1999) A Bayesian approach to sensitivity analysis. Health Econ, vol.8, pp. 263-268.

Fenwick E, Claxton K, Sculpher M, Briggs A. (2000) Improving the efficiency and relevance of health technology assessment: The role of iterative decision analytic modelling. CHE Discussion Paper 179. York, University of York.

Fenwick E, Claxton K, Sculpher M. (2001) Representing uncertainty: the role of cost-effectiveness acceptability curves. Health Econ, vol. 10, pp. 779-87.

Fleisch H. (1993) Bisphosphonates in osteoporosis: an introduction. Osteoporos Int, vol. 3, pp. S3- S5.

Francis RM, Anderson FH, Torgerson DJ. (1995) A comparison of the effectiveness and cost of treatment for vertebral fractures in women. Br J Rheumatol, vol. 34, pp. 1167-71.

Fussell J. (1975) How to hand calculate system reliability characteristics. IEEE Transactions on Reliability, 24.

Gabriel SE, Kneeland TS, Melton LJ, et al. (1999) Health-related quality of life in economic evaluations for osteoporosis: whose values should we use? Med Decis Making, vol. 19, pp. 141-8.

Garattini L, Grilli R, Scopelliti D, Mantovani L. (1995) A proposal for Italian guidelines in pharmacoeconomics. Pharmacoeconomics, vol. 7, pp. 1-6.

Glasziou PP, Mitchell AS. (1996)Use of pharmacoeconomic data by regulatory authorities. In: Spilker B (Ed). Quality of Life and Pharmacoeconomics in Clinical Trials. Philadelphia, Lippincott-Raven.

Goeree R, O'Brien B, Pettitt D, Cuddy L, Ferraz M, Adachi J. (1996) An assessment of the burden of illness due to osteoporosis in Canada. Journal SOGC, Supplement, pp.15-22.

Gold MR, Siegel JE, Russell LM, Weinstein MC. (1996) Cost-Effectiveness in Health and Medicine. New York, Oxford University Press.

Grima DT, Burge RT, Becker DL, Tosteson ANA. (2002) Short-term cost-effectiveness of bisphosphonate therapies for postmenopausal osteoporotic women at high risk of fracture. P&T, vol. 27, pp. 448-455.

Grimshaw GM, Szczepura A, Hulten M, MacDonald F, Nevin NC, Sutton F, Dhanjal S. (2003) Evaluation of molecular tests for prenatal diagnosis of chromosome abnormalities. Health Technol Assess, vol. 7, no. 10.

Hadorn DC. (1992) The problem of discrimination in health care priority setting. JAMA, vol. 268, pp. 1454-9

Hamby DM, Tarantola S. (1999) Exploring sensitivity analysis techniques for the assessment of an environmental transport model. In: Proceedings of ESREL '99 - The Tenth European Conference on Safety and Reliability. Garching, Germany.

Hamby DM. (1995) A comparison of sensitivity analysis techniques. Health Physics, vol. 68, pp. 195-204.

Harris ST, Watts NB, Genant HK et al. (1999) Effects of risedronate treatment on vertebral and nonvertebral fractures in women with postmenopausal osteoporosis: a randomized controlled trial. Vertebral Efficacy With Risedronate Therapy (VERT) Study Group. JAMA, vol. 282, pp. 1344-52.

Helton JC. (1993) Uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive waste disposal. Reliab Eng and Sys Safety, vol. 42, pp. 327-367.

Hertwich EG, McKone TE, Pease WS. (1999) Parameter uncertainty and variability in evaluative fate and exposure models. Risk Analysis, vol. 19, pp. 1193-1204.

Hillman AL, Eisenberg JM, Pauly MV et al. (1991) Avoiding bias in the conduct and reporting of cost effectiveness research sponsored by pharmaceutical companies. N Eng J Med, vol. 324, 1362-5.

Hoch JS, Briggs AH, Willan AR. (2002) Something old, something new, something borrowed, something blue: a framework for the marriage of health econometrics and cost-effectiveness analysis. Health Econ, vol. 11, pp. 415-30.

Hofer E. (1999) Sensitivity analysis in the context of uncertainty analysis for computationally intensive models. Computer Physics Communication, vol. 17, pp. 21-34.

Hoffman FO, Hammonds JS. (1994) Propagation of uncertainty in risk assessments: the need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability. Risk Analysis. vol. 14, pp.707-712

Hu R, Mustard CA, Burns C. (1996) Epidemiology of incident spinal fracture in a complete population. Spine, vol. 21, pp. 492-9.

Huijbregts MAJ, Thissen U, Jager T, van de Meent D, Ragas AMJ. (2000) Priority assessment of toxic substances in life cycle assessment. Part II: assessing parameter uncertainty and human variability in the calculation of toxicity potentials. Chemosphere, vol. 4, pp. 575-588.

Hummel S, Paisley S, Morgan A, Currie E, Brewer N. (2003) Clinical and cost-effectiveness of new and emerging technologies for early localised prostate cancer: a systematic review. Health Technol Assess, vol. 7, no. 33.

Hunink MG, Bult JR, de Vries J, Weinstein MC. (1998) Uncertainty in decision models analyzing cost-effectiveness: the joint distribution of incremental costs and effectiveness evaluated with a nonparametric bootstrap method. Med Decis Making, vol. 18, pp. 337-46.

Hutubessy RC, Baltussen RM, Evans DB, Barendregt JJ, Murray CJ. (2001) Stochastic league tables: communicating cost-effectiveness results to decision-makers. Health Econ, vol 10, pp 473-7.

Iglesias CP, Torgerson DJ, Bearne A, Bose U. (2002) The cost utility of bisphosphonate treatment in established osteoporosis. QJM, vol. 95, pp. 305-11.

Iman RL, Helton JC. (1988) An investigation of uncertainty and sensitivity analysis techniques for computer models. Risk Anal, vol. 8, pp. 71-90.

Jacobsen SJ, Goldberg J, Miles TP, Brody JA, Stiers W, Rimm AA. (1990) Regional variation in the incidence of hip fracture: US white women aged 65 years and older. JAMA, vol. 264, pp. 500-502.

James BR, Gwo JP, Toran LE. (1996) Risk-cost decision framework for aquifer remediation design. J Water Resour Plan and Manag, vol.122, pp. 414-420.

Johnell O, Gullberg B, Allander JA, KanisJA . (1992) The apparent incidence of hip fracture in Europe: A study of national register sources. Osteoporos Int, vol. 2, pp. 298-302.

Johnell O, Jonsson B, Jonsson L, Black D. (2003) Cost effectiveness of alendronate (Fosamax©) for the treatment of osteoporosis and prevention of fractures. Pharmacoeconomics, vol. 21, pp. 305-14.

Jones L, Hawkins N, Westwood M, Wright K, Richardson G, Riemsma R. Systematic review of the clinical effectiveness and cost-effectiveness of capecitabine (Xeloda) for locally advanced and/or metastatic breast cancer. Health Technol Assess, vol. 8, no. 5.

Jönsson L, Borgström F, Zethraeus N. (2002) Cost-Effectiveness of Alendronate in the Treatment of Osteoporosis in Denmark – An Economic Evaluation Based on the Fracture Intervention Trial. SSE/EFI Working Paper Series in Economics and Finance, No 501.

Kado DM, Browner WS, Palermo L, Nevitt MC, Genant HK, Cummings SR. (1999) Vertebral fractures and mortality in older women: a prospective study. Study of Osteoporotic Fractures Research Group. Arch Intern Med, vol. 159, pp. 1215-20.

Kaltenthaler E, Vergel YB, Chilcott J, Thomas S, Blakeborough T, Walters SJ, Bouchier H. (2004) A systematic review and economic evaluation of magnetic resonance cholangiopancreatography compared with diagnostic endoscopic retrograde cholangiopancreatography. Health Technol Assess, vol. 8, no. 10.

Kanis JA, Brazier JE, Stevenson M, Calvert NW, Lloyd Jones M. (2002) Treatment of established osteoporosis: a systematic review and cost-utility analysis. Health Technol Assess, vol. 6(29), pp. 1-146

Kanis JA, Melton LJI, Christiansen C, et al. (1994) The diagnosis of osteoporosis. J Bone Miner Res, vol. 9, pp. 1137-41.

Kapur JN (1989) Maximum entropy models in science and engineering. Wiley, New York.

Kassirer JP, Angell M. (1994) The Journals policy on cost-effectiveness analyses. N Eng J Med, vol. 331, pp. 669-70.

Klotzbuecher CM, Ross PD, Landsman PB, et al. (2000) Patients with prior fractures have an increased risk of future fractures: A summary of the literature and statistical synthesis. J Bone Min Res, vol. 15, pp. 721-739.

Krzykacz-Hausmann B. (2001) Epistemic sensitivity analysis based on the concept of entropy. In Proceedings of the Third International Symposium on Sensitivity Analysis of Model Output, Madrid, Spain.

Kuntz K, Weinstein M. (2001) Modelling in economic evaluation. In Drummond M, McGuire A (eds) Economic Evaluation in Health Care: Merging Theory with Practice. Oxford, Oxford University Press.

Kuntz KM, Goldie SJ. (2002) Assessing the sensitivity of decision-analytic results to unobserved markers of risk: defining the effects of heterogeneity bias. Med Decis Making, vol. 22, pp. 218-27.

Lang AE, Lozano AM. (1998) Parkinson's disease: Part One. N Eng J Med, vol. 339, pp. 1044-53.

Laupacis A. (2002) Inclusion of drugs in provincial drug benefit programs: Who is making these decisions, and are they the right ones? CMAJ, vol. 166, pp. 44-7.

Leibson CL, Tosteson ANA, Gabriel SE, Ransom JE, Melton LJ. (2002) Mortality, disability and nursing home use for persons with and without hip fracture: A population-based study. J Am Geriatr Soc, vol. 50, pp. 1644-1650.

Levitin G, Lisnianski A. (1999) Importance and sensitivity analysis of multi-state systems using universal generating functions method. Reliab Eng and Sys Safety, vol. 65, pp. 271-282.

Liberman UA, Weiss SR, Broll J, Minne HW, Quan H, Bell NH et al. (1995) Effect of oral alendronate on bone mineral density and the incidence of fractures in postmenopausal osteoporosis. The Alendronate Phase III Osteoporosis Treatment Study Group. N Eng J Med, vol. 333, pp. 1437-43.

Lipsey RG, Courant PN, Ragan CTS. (1999) Economics 12th ed. Addison-Wesley, Reading, Massachusetts.

Lothgren M, Zethraeus N. (2000) Definition, interpretation and calculation of cost-effectiveness acceptability curves. Health Econ, vol. 9, pp. 623-30.

Lyritis GP, Tsakalakos N, Paspati I, et al. (1997) The effect of a modified etidronate cyclical regimen on postmenopausal osteoporosis: a four-year study. Clin Rheumatol, vol. 16, pp. 354–360.

Maetzel A, Tugwell P, Boers M, Guillemin F, Coyle D, Drummond M, Wong JB, Gabriel SE. (2003) Economic evaluation of programs or interventions in the management of rheumatoid arthritis: defining a consensus-based reference case. J Rheum, vol. 30, pp. 891-6.

Magnusson SE, Frantzich H, Harada K. (1996) Fire safety design based on calculations. Fire Safety Journal, vol. 27, pp. 305-334.

Mark DB, Hlatky MA, Califf RM et al. (1995) Cost effectiveness of thrombolytic therapy with tissue plasminogen activator as compared to streptokinase for acute myocardial infarction. N Eng J Med, vol. 332, pp. 1418-24.

Marshall D, Johnell O, Wedel H. (1996) Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures. BMJ, vol. 312, pp. 1254-1259.

McClung MR, Geusens P, Miller PD et al. (2001) Effect of risedronate on the risk of hip fracture in elderly women. Hip Intervention Program Study Group. N Eng J Med, vol. 344, pp. 333-40.

McKenna SA, Arnold BW. (1998) Quantitative sensitivity and importance analysis of parameters in Monte-Carlo methods. In Proceedings of the International High-Level Radioactive Waste Management Conference, American Nuclear Society, La Grange Park, Illinois.

Meads C, Salas C, Roberts T, Moore D, Fry-Smith A, Hyde C. (2003) Clinical effectiveness and cost-utility of photodynamic therapy for wet age-related macular degeneration: a systematic review and economic evaluation. Health Technol Assess, vol. 7, no. 9.

Melton LJ, Thamer M, Ray NF et al. (1997) Fractures attributable to osteoporosis: report from the National Osteoporosis Foundation. J Bone Miner Res, vol. 12, pp. 16-23.

Meltzer D. (2001) Addressing uncertainty in medical cost-effectiveness analysis implications of expected utility maximization for methods to perform sensitivity analysis and the use of cost-effectiveness analysis to set priorities for medical research. J Health Econ, vol. 20, pp. 109-29.

Morgan MG, Henrion M. (1990) Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis. Cambridge University Press, New York.

Mortensen L, Charles P, Bekker PJ, Digennaro J, Johnston CC. (1998) Risedronate increases bone mass in an early postmenopausal population: two years of treatment plus one year of follow-up. J Clin Endocrinol Metab, vol. 83, pp. 396-402.

Mowatt G, Vale L, Perez J, Wyness L, Fraser C, MacLeod A, Daly C, Stearns SC. (2003) Systematic review of the effectiveness and cost-effectiveness, and economic evaluation, of home versus hospital or satellite unit haemodialysis for people with end-stage renal failure. Health Technol Assess, vol. 7, no. 2.

National Institute for Clinical Excellence. (2001) Guidance for Manufacturers and Sponsors. London: National Institute for Clinical Excellence.

NICE. (2004) Guide to the methods of technology appraisal. National Institute for Clinical Excellence, London.

Nuijten MJ, Hardens M. (1997) Measuring sensitivity in pharmacoeconomic studies. An integration of point-sensitivity and range-sensitivity. Pharmacoeconomics, vol. 12, pp. 555-64.

Nuijten MJ. (1999) Measuring sensitivity in pharmacoeconomic studies. Refining point sensitivity and range sensitivity by incorporating probability distributions. Pharmacoeconomics, vol. 16, pp. 33-41.

O'Brien BJ, Sculpher MJ. (2000) Building uncertainty into cost-effectiveness rankings: portfolio risk-return tradeoffs and implications for decision rules. Med Care 38:460-8.

O'Hagan A, Luce BR. (2003) A Primer on Bayesian Statistics in Health Economics. Centre for Bayesian Statistics in Health Economics, Sheffield.

O'Hagan A, Stevens JW. (2002) The probability of cost-effectiveness. BMC Med Res Methodol, vol.2, p. 5.

Ontario Case Costing Initiative. (2000) Ontario Guide to Case Costing (version 2). Toronto, Ontario Case Costing Initiative.

Ontario Ministry of Health and Long Term Care. (2000) Ontario Guidelines for Drug Submission and Evaluation. Toronto, Government of Ontario.

Ontario Ministry of Health and Long-term Care. (2001) Ontario Health Insurance (OHIP) Schedule of Benefit and Fees. Toronto, Government of Ontario.

Ontario Ministry of Health and Long Term Care. (2002) Ontario Drug Benefit Formulary/Comparative Drug Index No. 38. Toronto, Government of Ontario.

Ontario Ministry of Health and Long Term Care. (2003) Ontario Drug Benefit Formulary/Comparative Drug Index No. 38 Update A. Toronto, Government of Ontario.

Ontario Ministry of Health and Long Term Care. (2004) Handbook of Limited Use Drug Products. Toronto, Government of Ontario.

Palmer CS, Schmier JK, Snyder E, Scott B. (2000) Patient preferences and utilities for 'off-time' outcomes in the treatment of Parkinson's disease. Quality of Life Research, vol. 9, pp. 819-27.

Pandor A, Eastham J, Beverley C, Chilcott J, Paisley S. (2004) Clinical effectiveness and cost-effectiveness of neonatal screening for inborn errors of metabolism using tandem mass spectrometry: a systematic review. Health Technol Assess, vol. 8, no. 12.

Papadimitropoulos EA, Coyte PC, Josse RG, Greenwood CE. (1997) Current and projected rates of hip fracture in Canada. CMAJ, vol. 157, pp. 1357-63.

Papadimitropoulos M. (2000) PhD thesis, University of Toronto.

Pols HA, Felsenberg D, Hanley DA et al. (1999) Multinational, placebo-controlled, randomized trial of the effects of alendronate on bone density and fracture risk in postmenopausal women with low bone mass: results of the FOSIT study. Foxamax International Trial Study Group. Osteoporos Int, vol. 9, pp. 461-8.

Rawlins M. (2004) Completed technology appraisals 2000/2001. Report from the Chairman to the Board of NICE July 18th 2001. Accessed January 26th 2004, from http://www.nice.org.uk/pdf/brdjul01item3.pdf

Reginster J, Minne HW, Sorensen OH et al. (2000) Randomized trial of the effects of risedronate on vertebral fractures in women with established postmenopausal osteoporosis. Vertebral Efficacy with Risedronate Therapy (VERT) Study Group. Osteoporos Int, vol. 11, pp. 83-91.

Rivest J, Barclay CL, Suchowersky O. (1999) COMT inhibitors in Parkinson's disease. Can J Neurol Sci, vol. 26(Suppl 2), pp. S34-S38.

Roderick P, Davies R, Raftery J, Crabbe D, Pearce R, Bhandari P, Patel P. (2003) The cost-effectiveness of screening for Helicobacter pylori to reduce mortality and morbidity from gastric cancer and peptic ulcer disease: a discrete-event simulation model. Health Technol Assess, vol. 7, no. 6.

Rosner AJ, Grima DT, Torrance GW, Bradley C, Adachi JD, Sebaldt RJ, Willison DJ. (1998) Cost effectiveness of multi-therapy treatment strategies in the prevention of vertebral fractures in postmenopausal women with osteoporosis. Pharmacoeconomics, vol. 14, pp. 559-73.

Ross JR, Saunders Y, Edmonds PM, Patel S, Wonderling D, Normand C, Broadley K. (2004) A systematic review of the role of bisphosphonates in metastatic disease. Health Technol Assess, vol. 8, no. 4.

Rovira J, Antoñanzas F. (1995) Economic analysis of health technologies and programmes. A Spanish proposal for methodological standardization. Pharmacoeconomics, vol. 8, pp. 245-252.

Saltelli A, Chan K, Scott EM. (2000) Sensitivity Analysis. Wiley, Chichester.

Sculpher M, Torgerson D, Goeree R, O'Brien B. (1999) A critical structured review of economic evaluations of interventions for the prevention and treatment of osteoporosis. Centre for Health Economics Discussion Paper 169. York, Univeristy of York.

Sharp PF, Olson J, Strachan F, Hipwell J, Ludbrook A, O'Donnell M, Wallace S, Goatman K, Grant A, Waugh N, McHardy K, Forrester JV. (2003) The value of digital imaging in diabetic retinopathy. Health Technol Assess, vol. 7, no. 30.

Song FJ, Barton P, Sleightholme V, Yao GL, Fry-Smith A. (2003) Screening for fragile X syndrome: a literature review and modelling study. Health Technol Assess, vol. 7, no. 16.

Spear RC, Hornberger GM. (1980) Eutrophication in Peel Inlet, II. Identification of critical uncertainties via generalized sensitivity analysis. Water Research, vol. 14, pp. 43-49.

Spiegelhalter DJ, Myles JP, Jones DR, Abrams KR. (2000) Bayesian methods in health technology assessment: a review. Health Technol Assess, vol. 4(38), pp. 1-130.

Statistics Canada. (1995a) Life Tables, Canada and Provinces, 1990-92. 84-537 Occasional. Ottawa, Statistics Canada.

Statistics Canada. (1995b) National Population Health Survey. Ottawa: Statistics Canada.

Steen H, Erikstad KE. (1996) Sensitivity of willow grouse Lagopus lagopus population dynamics to variations in demographic parameters. Wildlife Biology, vol. 2, pp. 27-35.

Stinnett AA, Mullahy J (1998). Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. Med Decis Making, vol. 18, pp. S68-S80.

Stinnett AA. Paltiel AD. (1997) Estimating CE ratios under second-order uncertainty - the mean ratio versus the ratio of means. Med Decis Making, vol. 17, pp. 483-489.

Storm T, Thamsborg G, Steiniche T, Genant HK, Sorensen OH. Effect of intermittent cyclical etidronate therapy on bone mass and fracture rate in women with postmenopausal osteoporosis. N Engl J Med, vol. 322, pp. 1265-71.

Tenenhouse A, Joseph L, Kreiger N et al. (2000) Estimation of the prevalence of low bone density in Canadian women and men using a population-specific DXa reference standard: The Canadian Multicentre Osteoporosis Study (CaMos). Osteoporos Int, vol. 10, pp. 897-904.

Thompson KM, Graham JD. (1996) Going beyond the single number: using probabilistic risk assessment to improve risk management. Human and Ecological Risk Assessment, vol. 2, pp. 1008-1034.

Tonino RP, Meunier PJ, Emkey R, et al. (2000) Skeletal benefits of alendronate: 7-year treatment of postmenopausal osteoporotic women. J Clin Endocrinol Metab, vol. 85, pp. 3109-3115.

Torgerson DJ, Kanis JA. (1995) Cost-effectiveness of preventing hip fractures in the elderly population using vitamin D and calcium. QJM, vol. 88, pp. 135-9.

Torrance GW, Feeny D. (1989) Utilities and quality-adjusted life years. Int J Technol Assess Health Care, vol. 5, pp. 559-575.

Torrance GW, Furlong W, Feeny D, Boyle M. (1995) Multi-attribute preference functions. Health Utilities Index. Pharmacoeconomics, vol. 7, pp. 503-20.

Tosteson AN, Grove MR, Hammond CS, Moncur MM, Ray GT, Hebert GM, Pressman AR, Ettinger B. (2003) Early discontinuation of treatment for osteoporosis. Am J Med, vol. 115, pp. 209-16.

Tosteson AN, Jonsson B, Grima DT, O'Brien BJ, Black DM, Adachi JD. (2001a) Challenges for model-based economic evaluations of postmenopausal osteoporosis interventions. Osteoporos Int, vol. 12, pp. 849-57.

Tosteson ANA, Gabriel SE, Grove MR et al. (2001b) Impact of hip and vertebral fractures on quality adjusted life years. Osteoporos Int, vol. 12, pp. 1042-1049.

Tosteson ANA, Hammond CS. (2002) Quality-of-life assessment in osteoporosis: health-status and preference-based measures. Pharmacoeconomics, vol. 20, pp. 289-303

Turner D, Wailoo A, Nicholson K, Cooper N, Sutton A, Abrams K. (2003) Systematic review and economic decision modelling for the prevention and treatment of influenza A and B. Health Technol Assess, vol. 7, no. 35.

van Hout BA, Al MJ, Gordon GS, Rutten FF. (1994) Costs, effects and C/E-ratios alongside a clinical trial. Health Econ, vol. 3, pp. 309-19.

Visentin P, Ciravegna R, Fabris F. (1997) Estimating the cost per avoided hip fracture by osteoporosis treatment in Italy. Maturitas, vol. 26, pp. 185-92.

Waldegger L, Cranney A, Man-Son-Hing M, Coyle D. (2003) Cost-effectiveness of hip protectors in institutional dwelling elderly. Osteoporos Int, vol. 14, pp. 243-50.

Wardlaw JM, Keir SL, Seymour J, Lewis S, Sandercock PA, Dennis MS, Cairns J. (2004) What is the best imaging strategy for acute stroke? Health Technol Assess, vol. 8, no. 1.

Watts NB, Harris ST, Genant HK et al. (1990) Intermittent cyclical etidronate treatment of postmenopausal osteoporosis. N Engl J Med, vol. 323, pp. 73-9.

Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C, Luce BR. (2003) Principles of good practice for decision analytic modelling in health care evaluation: Report of the ISPOR Task Force on Good Research Practices-Modeling Studies. Value in Health, vol. 6, pp. 9-17.

WHO Collaborating Centre for Public Health Aspects of Osteoporosis and other Rheumatic Diseases. (1999) Recommendations for Health Economics Evaluations of Interventions in Osteoporosis. Switzerland, World Health Organization.

WHO study group. (1994) Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. Switzerland, World Health Organization.

Wiktorowicz ME, Goeree R, Papaioannou A, Adachi JD, Papadimitropoulos E. (2001) Economic implications of hip fracture: health service use, institutional care and cost in Canada. Osteoporos Int, vol. 12, pp. 271-8.

Willan AR, O'Brien BJ. (1996) Confidence intervals for cost-effectiveness ratios: an application of Fieller's theorem. Health Econ, vol. 5, pp. 297-305.

Williams A. (1997) Intergenerational equity: An exploration of the fair innings argument. Health Econ, vol. 6, pp. 117-132

Willis MS. (2002) The health economics of calcium and vitamin D3 for the prevention of osteoporotic hip fractures in Sweden. Int J Technol Assess Health Care, vol. 18, pp. 791-807.

Wimalawansa SJ. (1998) A four-year randomized controlled trial of hormone replacement and bisphosphonate, alone or in combination in women with postmenopausal osteoporosis. Am J Med, vol. 104, pp. 219-26.

Zaric GS. (2003) The impact of ignoring population heterogeneity when Markov models are used in cost-effectiveness analysis. Med Decis Making, vol. 23, pp. 379-96.

Zeikenfondsraad. (1999) Dutch Guidelines for Farmacoeconomic Research. Amstelveen Zeikenfondsraad, Netherlands.

Zethraeus N, Johannesson M, Jonsson B, Lothgren M, Tambour M. (2003a) Advantages of using the net-benefit approach for analysing uncertainty in economic evaluation studies. Pharmacoeconomics, vol. 21, pp. 39-48.

Zethraeus N, Sendrine B, Caulin F, Corcaud S, Gathon HJ, Haim M, Johnell O, Jonsson B, Kanis JA, Tsouderos Y, Reginster JY. (2003b) Models for assessing the cost-effectiveness of the treatment and prevention of osteoporosis. Osteoporos Int, vol. 13, pp. 841-857.