

Richard Edlin · Christopher McCabe  
Claire Hulme · Peter Hall  
Judy Wright

# Cost Effectiveness Modelling for Health Technology Assessment

A Practical Course



# Cost Effectiveness Modelling for Health Technology Assessment



Richard Edlin • Christopher McCabe  
Claire Hulme • Peter Hall • Judy Wright

# Cost Effectiveness Modelling for Health Technology Assessment

A Practical Course



Richard Edlin  
Faculty of Medicine and Health Sciences  
University of Auckland  
Auckland  
New Zealand

Christopher McCabe  
Department of Emergency Medicine  
University of Alberta  
Edmonton  
Alberta  
Canada

Claire Hulme  
Faculty of Medicine and Health  
University of Leeds  
Leeds  
West Yorkshire  
UK

Peter Hall  
Edinburgh Cancer Research Centre  
University of Edinburgh  
Edinburgh  
Scotland  
UK

Judy Wright  
Faculty of Medicine and Health  
University of Leeds  
Leeds  
West Yorkshire  
UK

ISBN 978-3-319-15743-6      ISBN 978-3-319-15744-3 (eBook)  
DOI 10.1007/978-3-319-15744-3

Library of Congress Control Number: 2015941443

Springer Cham Heidelberg New York Dordrecht London  
© Springer International Publishing Switzerland 2015

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

Adis is a brand of Springer  
Springer International Publishing AG Switzerland is part of Springer Science+Business Media  
([www.springer.com](http://www.springer.com))

# Preface

Cost effectiveness analysis for health interventions has come a long way over the last couple of decades. The methods and statistical techniques and the advent of modelling used in this context have evolved exponentially. For the analyst, this has meant a steep learning curve and the need to develop skills and expertise in decision-analytic modelling steeped in Bayesian methodology. The language that surrounds cost effectiveness analysis or cost effectiveness modelling has also evolved and can hinder understanding of these methods; especially given phrases like *economic evaluation*, *cost effectiveness* and *cost benefit* are used in everyday life – often interchangeably. This has become more and more apparent to us over the last 10 years as we have taught students in the methods of economic evaluation and cost effectiveness modelling.

So how does cost effectiveness modelling fit within cost effectiveness analysis? Put simply, cost effectiveness modelling (also known as decision-analytic cost effectiveness modelling) is often referred to as a *vehicle* for cost effectiveness analysis. An easy way to think about this is that economic evaluation is a range of methods that may be used to assess the costs and benefits (e.g. cost effectiveness analysis or cost–benefit analysis), and cost effectiveness analysis is one of those methods of economic evaluation. A decision-analytic model is a statistical method to inform decision processes and thus, a (decision-analytic) cost effectiveness model is a statistical method used to inform a decision process that incorporates cost effectiveness analysis. It is these cost effectiveness models that are the focus of this book.

Given the complex nature of cost effectiveness modelling and the often unfamiliar language that runs alongside it, we wanted to make this book as accessible as possible whilst still providing a comprehensive, in-depth, practical guide that reflects the state of the art – that includes the most recent developments in cost effectiveness modelling. Although the nature of cost effectiveness modelling means that some parts are inevitably ‘techy’, we have broken down explanations of theory and methods into bite-sized pieces that you can work through at your own pace, we have provided explanations of terms and methods as we use them and importantly, the exercises and online workbooks allow you to test your skills and understanding as you go along. The content and the exercises in the text have in large part been honed

by our students, particularly those who have attended the modelling courses we developed and run at the University of Alberta. A big thank you to those students!

Auckland, New Zealand  
Edmonton, AB, Canada  
Leeds, UK  
Edinburgh, UK  
Leeds, UK

Richard Edlin, PhD  
Christopher McCabe, PhD  
Claire Hulme, PhD  
Peter Hall, MBChB, PhD  
Judy Wright, MSc

# Acknowledgements

Christopher McCabe's academic programme is funded by the Capital Health Research Chair Endowment at the University of Alberta. Additional funding for work that contributed to the development of the material in this book was received from Genome Canada, Alberta Innovates Health Solutions, Canadian Institutes of Health Research, the Stem Cell Network and the UK National Institute for Health Services Research.

The authors would like to thank Klemens Wallner, Mike Paulden, Maria Sandova and Mira Singh (University of Alberta) and Alison Smith and David Meads (University of Leeds) for their help in the development of the material in this book. We particularly thank the investigators on the NIHR OPTIMA-prelim study for allowing us to use some of the study data for examples included in the book. Thanks are also due to the participants in the 2013 and 2014 Introduction to Cost effectiveness modelling courses held at the University of Alberta for their helpful comments on earlier versions of some of the course material.

November 2014

Richard Edlin  
Christopher McCabe  
Claire Hulme  
Peter Hall  
Judy Wright



# Contents

<b>1 Economic Evaluation, Cost Effectiveness Analysis and Health Care Resource Allocation .....</b>	1
1.1 Introduction.....	1
1.2 Scarcity, Choice and Opportunity Cost .....	2
1.3 Types of Economic Evaluation.....	2
1.3.1 Cost Benefit Analysis (CBA) .....	3
1.3.2 Cost Effectiveness Analysis (CEA) .....	3
1.3.3 Cost Utility Analysis (CUA). ....	5
1.4 Incremental Cost Effectiveness Ratios (ICERs).....	6
1.4.1 Simple and Extended Dominance.....	9
1.4.2 The Net Benefit Approach .....	11
1.5 Summary.....	12
References .....	13
<b>2 Finding the Evidence for Decision Analytic Cost Effectiveness Models .....</b>	15
2.1 Introduction.....	15
2.2 Choosing Resources to Search for Evidence .....	17
2.3 Designing Search Strategies .....	18
2.4 Searching for Existing Cost Effectiveness Models.....	19
2.4.1 Where to Look .....	20
2.4.2 Search Strategy, Concepts, Terms and Combinations .....	21
2.4.3 Search Filters, Database Limits and Clinical Queries .....	22
2.5 Searching for Clinical Evidence.....	23
2.5.1 Finding the Evidence on Incidence, Prevalence and Natural History of a Disease .....	23
2.5.2 Finding the Evidence on the Clinical Effectiveness of Health Interventions .....	24
2.5.3 Database Limits and Clinical Queries .....	28

2.6	Finding the Evidence on Health-Related Quality of Life and Health State Preferences .....	30
2.6.1	Where to Look .....	30
2.6.2	Search Strategy, Concepts, Terms and Combinations .....	31
2.6.3	Search Filters, Database Limits and Clinical Queries .....	32
2.7	Finding Evidence on Resource Use and Costs .....	32
2.7.1	Where to Look .....	34
2.7.2	Search Strategy, Concepts, Terms and Combinations .....	34
2.7.3	Search Filters, Database Limits and Clinical Queries .....	36
2.8	Tracking and Reporting Search Activities .....	36
2.9	Quality Assessment Tools. ....	37
2.10	Summary.....	38
	References .....	38
<b>3</b>	<b>Building a Decision Tree Cost Effectiveness Model .....</b>	<b>41</b>
3.1	Introduction. ....	41
3.2	What Is a Decision Model?. ....	42
3.3	Key Elements of a Decision Tree .....	44
3.4	Costs, Benefits and Complexity .....	49
3.5	Exercise Building a Decision Tree .....	51
3.6	Summary.....	56
	References .....	57
<b>4</b>	<b>Uncertainty, Probabilistic Analysis and Outputs from Cost Effectiveness Analyses. ....</b>	<b>59</b>
4.1	Introduction. ....	59
4.2	Sources of Uncertainty in Cost Effectiveness Models .....	60
4.2.1	Sampling Variation .....	60
4.2.2	Extrapolation. ....	61
4.2.3	Generalisability. ....	62
4.2.4	Model Structure .....	62
4.2.5	Methodological Uncertainty.....	63
4.3	Analytic Responses to Uncertainty in CEA .....	64
4.3.1	One-Way Sensitivity Analysis .....	65
4.3.2	Multiway Sensitivity Analysis .....	66
4.3.3	Threshold Analysis.....	67
4.3.4	Analysis of Extremes .....	67
4.4	Probabilistic Sensitivity Analysis (PSA) .....	68
4.5	Outputs from Probabilistic Analysis. ....	71
4.6	Some Problems with ICERs .....	72
4.7	Summary.....	75
	References .....	76
<b>5</b>	<b>Introduction to Markov Cost Effectiveness Models .....</b>	<b>77</b>
5.1	Introduction. ....	77
5.2	Why Use Markov Models?.....	78
5.3	Health States .....	79

5.4	Transition Probabilities . . . . .	80
5.5	Markov Trace . . . . .	81
5.6	Cycle Length, Time Horizon and Discounting . . . . .	83
5.7	Summary . . . . .	86
	References . . . . .	86
<b>6</b>	<b>Probability Distributions for Effectiveness Parameters . . . . .</b>	<b>87</b>
6.1	Introduction . . . . .	87
6.2	What Do We Mean by Effectiveness Parameters? . . . . .	88
6.2.1	Obtaining Information on Effectiveness . . . . .	88
6.3	Choosing Distributions for Effectiveness Parameters . . . . .	89
6.3.1	Fitting a Distribution . . . . .	90
6.4	Beta Distribution for Probabilities . . . . .	91
6.5	Dirichlet Distribution for Multinomial Probabilities . . . . .	94
6.6	Normal Distribution for Log-Relative Risk . . . . .	95
6.7	Survival Analysis for Time-to-Event Data . . . . .	96
6.7.1	The Exponential Distribution . . . . .	97
6.7.2	The Weibull Distribution . . . . .	98
6.7.3	The Gompertz Distribution . . . . .	99
6.7.4	Choice of Distribution for Time-to-Event Data . . . . .	100
6.8	Parameter Correlation in Survival Analysis . . . . .	102
6.9	Summary . . . . .	103
	References . . . . .	103
<b>7</b>	<b>Probability Distributions for Cost and Utility Parameters . . . . .</b>	<b>105</b>
7.1	Introduction . . . . .	105
7.2	Distributions for Cost Parameters . . . . .	106
7.2.1	The LogNormal Distribution . . . . .	106
7.2.2	The Gamma Distribution . . . . .	108
7.3	Distributions for Utility Parameters . . . . .	109
7.3.1	Distributional Characteristics of the Utility Scale . . . . .	111
7.4	Characterising Uncertainty for Expected Utility Values Close to 1.0 . . . . .	112
7.4.1	Characterising Uncertainty for Expected Utility Values Away from 1.0 . . . . .	112
7.4.2	Logical Ordering for Utilities in Cost Effectiveness Models . . . . .	114
7.4.3	Health State-Specific Side Effect Utility Decrements . . . . .	116
7.5	Summary . . . . .	117
	References . . . . .	117
<b>8</b>	<b>Correlated Parameters and the Cholesky Decomposition . . . . .</b>	<b>119</b>
8.1	Introduction . . . . .	119
8.2	Correlated Parameters . . . . .	120
8.3	Defining a Set of Correlated Parameters . . . . .	121
8.4	The Cholesky Decomposition . . . . .	125

8.5	What If I Need a Cholesky Decomposition for a Different Number of Variables? . . . . .	128
8.6	Interpreting the Cholesky Decomposition . . . . .	128
8.7	Summary . . . . .	129
Appendix	Appendix 8.1: Extending the Cholesky Decomposition for More Than Three Correlated Parameters . . . . .	130
<b>9</b>	<b>Building a Markov Cost Effectiveness Model in Excel</b> . . . . .	133
9.1	Introduction . . . . .	133
9.2	The Model . . . . .	134
9.3	Modelling in Excel . . . . .	136
9.4	Constructing the Parameter Table . . . . .	138
9.5	Programming Your Model . . . . .	139
9.6	Adding a Discount Rate, Costs and Utilities . . . . .	141
9.7	Adding the Calculation of the Deterministic Incremental Cost Effectiveness Ratio (ICER) . . . . .	142
9.8	Summary . . . . .	143
<b>10</b>	<b>Making a Markov Model Probabilistic</b> . . . . .	145
10.1	Introduction . . . . .	145
10.2	Deterministic and Probabilistic Cost Effectiveness Analysis . . . . .	145
10.3	Making Model Parameters Stochastic . . . . .	147
10.4	Obtaining a Probabilistic Sensitivity Analysis from a Stochastic Model . . . . .	149
10.5	Exercise: Probabilistic Effectiveness Parameters . . . . .	151
10.6	Exercise: Probabilistic Cost and Utility Parameters . . . . .	153
10.7	Exercise: Incorporating the Cholesky Decomposition . . . . .	156
10.8	Summary . . . . .	159
Appendix	Appendix 10.1: Optimising Visual Basic Macros in Excel . . . . .	160
<b>11</b>	<b>Outputs from Probabilistic Sensitivity Analysis</b> . . . . .	163
11.1	Introduction . . . . .	163
11.2	Scatter Plots on the Cost Effectiveness Plane . . . . .	164
11.3	Cost Effectiveness Acceptability Curves (CEACs) . . . . .	167
11.4	Cost Effectiveness Acceptability Frontiers (CEAFs) . . . . .	169
11.5	Scatter Plots, CEACs and CEAF Exercises . . . . .	171
11.6	Summary . . . . .	174
References	References . . . . .	174
<b>12</b>	<b>Investing in Health Care, Research and the Value of Information</b> . . . . .	177
12.1	Introduction . . . . .	177
12.2	Uncertainty and Health-Care Reimbursement Decision-Making Processes . . . . .	178
12.3	Investing in Innovative Health Technologies . . . . .	179

12.4	Net Benefit Probability Map and Managing Decision Uncertainty .....	181
12.5	Delaying a Reimbursement Decision for More Research .....	185
12.5.1	Uncertainty in Decision Making and the Cost of Making the Wrong Decision .....	186
12.5.2	Expected Value of Perfect Information and the Value of Sample Information.....	187
12.5.3	Calculating the Expected Value of Perfect Information .....	190
12.6	Disaggregating the Value of Information: Expected Value of Perfect Parameter Information and the Expected Value of Sample Information .....	191
12.6.1	Expected Value of Perfect Parameter Information .....	191
12.6.2	Expected Value of Sample Information .....	192
12.7	Exercise: Constructing the Net Benefit Probability Map and Calculating the Value of Perfect Information .....	192
12.7.1	Calculating the Expected Value of Perfect Information .....	193
12.8	Summary.....	196
	References .....	197
<b>13</b>	<b>Value of Information in Health Technology Regulation and Reimbursement.....</b>	<b>199</b>
13.1	Introduction.....	199
13.2	Value of Information Analysis for Research Prioritisation.....	200
13.3	Value of Information Analysis for Research Design .....	200
13.3.1	Calculating the Expected Net Present Value of Sample Information .....	201
13.4	Is Decision Theory Ready to Inform Trial Design?.....	203
13.4.1	Structuring a Decision Problem .....	204
13.4.2	Evidence Synthesis and Model Parameterisation.....	204
13.4.3	Computational and Statistical Challenges .....	204
13.4.4	Adoption by Regulatory Organisations and Reimbursement Agencies .....	204
13.4.5	Adoption by Public Research Commissioners and Clinical Trialists.....	205
13.4.6	Industrial Development of Health Technologies .....	205
13.5	Value of Information in the Evolving Regulatory and Reimbursement Environments.....	206
13.6	Summary.....	206
	Appendix .....	207
	Appendix 13.1: General Monte Carlo Sampling Algorithm for Calculation of Population ENPVSI.....	207
	References .....	208

# **Chapter 1**

## **Economic Evaluation, Cost Effectiveness Analysis and Health Care Resource Allocation**

**Abstract** In the face of limited or scarce resources, how does the health care sector make decisions about what to prescribe or to recommend for patients; how do they decide which new technologies, programmes or service delivery models to adopt; and how do they decide what represents acceptable value for money? The purpose of this book is to provide an introduction to decision analytic cost effectiveness modelling, providing the theoretical and practical knowledge required to design and implement analyses to help answer these questions, models that meet the methodological standards of health technology assessment organisations. This introductory chapter provides an overview of economic evaluation, cost effectiveness analysis and health care resource allocation.

### **1.1 Introduction**

The purpose of this book is to provide an introduction to decision analytic cost effectiveness modelling, providing students with the theoretical and practical knowledge required to design and implement analyses that meet the methodological standards of health technology assessment organisations like the UK National Institute for Health and Clinical Excellence (NICE) and the Canadian Agency for Drugs and Technologies in Health Care (CADTH) (NICE 2013; CADTH 2006). The book guides you through building a decision tree and Markov model and, importantly, shows how the results of cost effectiveness analyses are interpreted. It provides exercises to allow you to put into practice what you have learnt. Before we begin this process, this chapter provides an overview of economic evaluation, cost effectiveness analysis and health care resource allocation. Section 1.2 introduces the concepts of scarcity, choice and opportunity cost. Section 1.3 summarises different types of economic evaluation with focus on cost effectiveness and cost utility analyses. Section 1.4 then goes on to introduce the concepts of incremental cost effectiveness ratios, dominance and net benefit.

## 1.2 Scarcity, Choice and Opportunity Cost

Economics is based on the premise of scarcity; that there are limited resources (e.g. the time of a surgeon, specialised equipment or number of beds in a ward) and unlimited wants (the unlimited needs of patients). In the face of limited or scarce resources, how does the health care sector make decisions about what to prescribe or to recommend for patients; how do they decide which new technologies, programmes or service delivery models to adopt; and how do they decide what represents acceptable value for money? When deciding, all health care sectors and decision makers are constrained by budgets. In public systems, the money devoted to health care is often allocated to departments or ministries with responsibility for health. Even private providers of health care will typically have a limited budget to spend on providing health care, as all expenditures must be financed. Economic evaluation facilitates comparisons between health care programmes, treatments, services and interventions in terms of both the costs and consequences of those interventions. Whilst results of such analyses do not provide a definitive answer to how resources should be allocated, they act as a tool for use in the decision-making process by identifying what might happen when resources are allocated in different ways.

It is important to note that any spending choices or decisions that are made about health care provision incur an opportunity cost. The concept of opportunity cost is fundamental to health economics. It is based upon the idea that scarcity means that use of resources on one health care activity inevitably means sacrificing activity somewhere else. The most important issue when deciding whether to provide a health care intervention, service or programme is the extent to which it improves health; but there is no such thing as a free lunch! The more a health care intervention or programme costs, the fewer resources are available for other programmes or interventions. However, economic evaluation is not a form of accountancy or about cutting cost; it is about evaluating services in terms of their benefits and costs to provide information to allocate resources efficiently. It is the comparative analysis of alternative courses of action in terms of both their costs and consequences (Drummond et al. 1987). Thus, economic evaluation aims to provide robust information to inform choices with the aim of ensuring the benefits of programmes that are implemented exceed their opportunity costs, to help target resources to the greatest effect (World Bank 1993).

## 1.3 Types of Economic Evaluation

Economic evaluations can take a number of forms, including cost effectiveness analysis, cost benefit analysis and cost utility analysis. The focus of this book lies on cost effectiveness and cost utility analysis because cost benefit analysis is less frequently used in resource allocation decisions in the health care sector. However,

for completeness a brief description of cost benefit is provided. It is also important to note that whilst we have presented cost effectiveness analysis and cost utility analysis separately in this chapter, the term cost effectiveness is used as a ‘catch all’ term for analyses. When used in this sense, cost utility analysis is thought of as a special type of cost effectiveness analysis.

### ***1.3.1 Cost Benefit Analysis (CBA)***

CBA compares the benefits with the costs of an intervention, where benefits are valued in monetary terms. The analysis can inform whether a particular goal is worthwhile and how much more or less of society’s resources should be allocated to achieving that particular goal.

The basis of CBA is the idea that social welfare exists, can be expressed and can be maximised by moving additional resources to aspects of production where there is greater social benefit. The CBA decision rule rests on the principle that health interventions should be provided only if the monetary value of additional benefits of providing the intervention exceeds the additional costs required to do so. By choosing these interventions they might be financed in a way that ensures everyone in society will be better off. If the intervention’s costs exceed their benefits, there is no way to finance the intervention without making someone worse off. However, this assumes that it is possible to separate one intervention from another, there is the possibility of choice between them, that it is possible to estimate the outcomes of each intervention, to value these outcomes in monetary terms and to estimate the cost of providing each intervention.

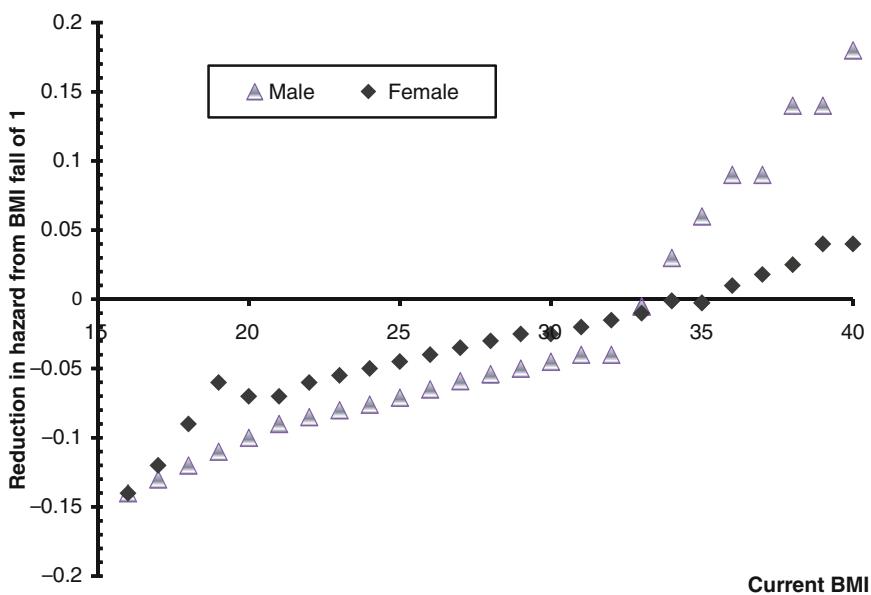
### ***1.3.2 Cost Effectiveness Analysis (CEA)***

CEA has its roots in production theory. The analysis attempts to identify where more health benefit can be produced at the same cost or where the same health benefit can be achieved for a lower cost. The social decision-making approach used in CEA is based on the premise that the aim of economic evaluation is to maximise whatever the decision maker wants to maximise (Sugden and Williams 1978). As such only those costs and benefits that the decision maker finds relevant need to be included in the analysis. The decision maker may be the society, the public sector, the health sector, the patient and their carer or family.

Thus the results of CEA depend in part on the perspective of the decision maker. In the health care field, this has often led to only health care system costs being included in CEA, with the argument that the health care budget should be used to maximise health (Johansson 1991). A fixed budget can be used to maximise the health effects using information about the incremental cost effectiveness ratios of different health care programmes or interventions. Alternatively, a price

per effectiveness unit can be set and used as the decision rule. In practice a single budget used to maximise the health effects must be identified to follow the budget maximisation approach (assuming the decision maker wants to maximise the health effects using the health care budget). However, only costs that fall on this budget will be included in the analyses. This can lead to suboptimal decisions from a societal perspective as costs outside the health system budget are ignored. Whilst bodies such as NICE provide guidance based on cost effectiveness from the health and social care budget perspective (NICE 2013), there has been movement towards a societal perspective that includes, for example, lost productivity (DoH 2010).

Unlike CBA, outcomes within CEA are measured in natural units (e.g. life years saved, cancers detected, reduction in blood pressure, heart attacks avoided). To be valid, the outcome needs to have a consistent value; the value shouldn't be dependent on the person (it should be comparable), and the value attached to each change in the outcome should only depend on how big that change is (it should have interval properties). Consider a hypothetical trial which uses body mass index (BMI) as the primary outcome. Suppose that this trial suggests that the risks are higher for those with higher and lower weight. Fig. 1.1 shows how the hazards ratios (the risks people face) differ when a patient's BMI falls by 1 point ( $\text{kg}/\text{m}^2$ ). If the comparability and interval properties held for BMI, we would expect a constant figure for all potential patients, and across all values of current BMI. It is clear from the diagram that this



**Fig. 1.1** BMI comparability and interval property

is not the case, since the value of a BMI reduction increases as current BMI increases, and is not the same across males and females. As a result, it would appear that BMI reduction would not be an appropriate outcome for cost effectiveness analysis.

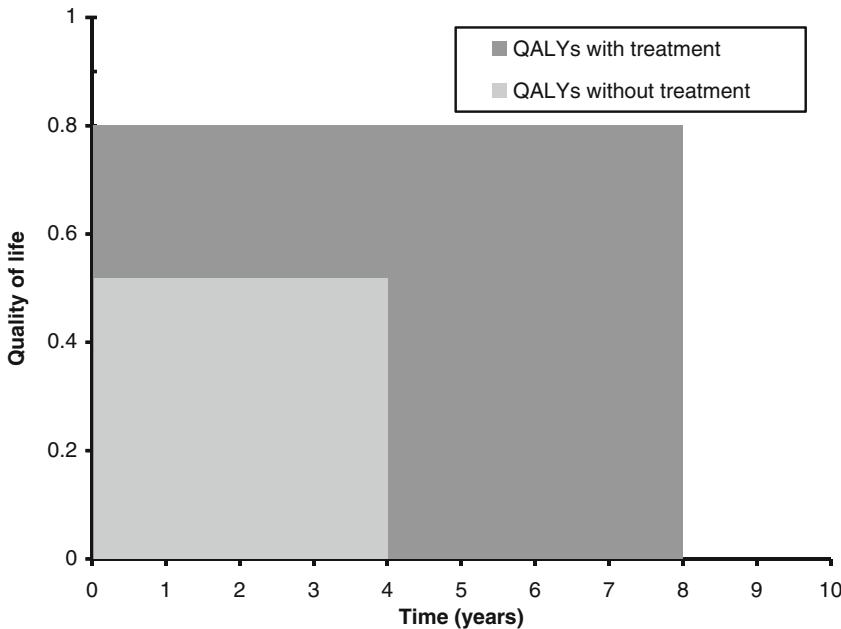
CEA are relatively simple to carry out. However, because the analysis only considers one outcome, it cannot incorporate other aspects of effect into the cost effectiveness ratio; hence, a CEA that reports the cost per life year saved will not capture potentially important impacts on patients' quality of life. In addition interventions with different objectives cannot be compared. How does a decision maker compare the cost per heart attack avoided and the cost per hip fracture avoided? A standard CEA will not help. A further weakness is that the relationship between outcome measure and health is often unclear, especially when the unit of effect is a biological measure such as tumour response or prostate-specific antigen reading.

In an economic evaluation, both the difference in cost between two options (also called the 'incremental cost') and the difference in effectiveness ('incremental effectiveness') are important. We cannot be certain in any analysis as to exactly how much either costs or effectiveness differs. Cost minimisation analysis (CMA) is a special type of analysis which assumes that we can be certain, at least as far as effectiveness is concerned. Under CMA, the analysis assumes that there is no difference in effectiveness and considers only the costs associated with each intervention. Not only is the founding assumption indefensible, it has been shown that adopting this assumption produces biased results as it ignores the correlation between magnitude of effect and cost. As the assumption underlying CMA is thought to be unhelpful, this type of analysis has largely fallen out of favour (Brazier et al. 2007; Dakin and Wordsworth 2013).

### 1.3.3 Cost Utility Analysis (CUA)

CUA is generally regarded as a more sophisticated version of CEA – and often just called CEA. The two analyses differ in how outcomes are measured. Within CUA the effect is measured in terms of 'healthy years'. Healthy years are represented by a multidimensional utility-based measure which combines life years gained with some judgement on the quality of those life years. Measures include quality adjusted life years (QALYs) and disability adjusted life years (DALYs). Both are used in CUA, but since the 1990s the QALY has been widely accepted as the reference standard in CUA (Gold et al. 1996; NICE 2013).

In the previous paragraph we referred to a *utility-based measure*. A utility is a measure of preference. In this context a utility is the measure of the preference or value that an individual or society places upon a particular health state. Usually the maximum value, 1 = full health/perfect health and 0 = dead. Health states that are considered worse than dead take a negative value. These utility values can be combined with survival data to derive QALYs. Consider the following example. An individual has a health condition with a utility value of 0.5 and current life expectancy of 4 years at a constant health state (i.e. their health state doesn't deteriorate



**Fig. 1.2** Illustration of the QALY gain

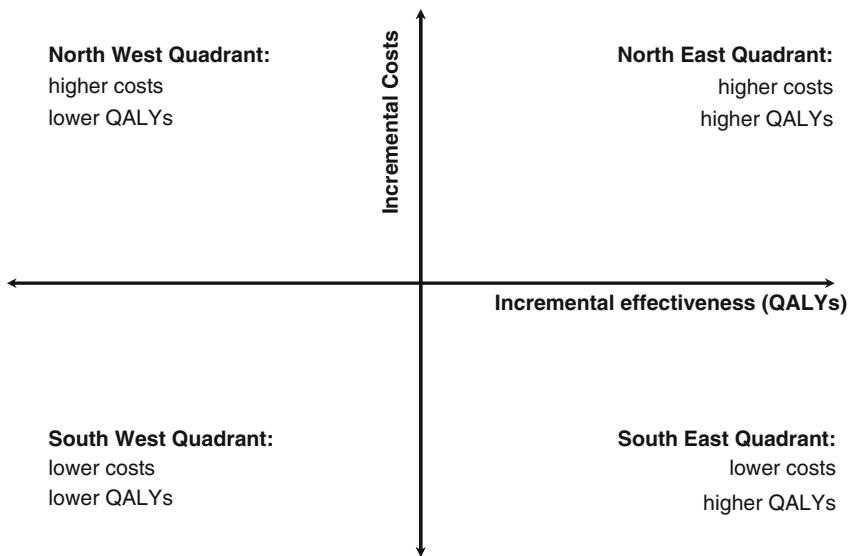
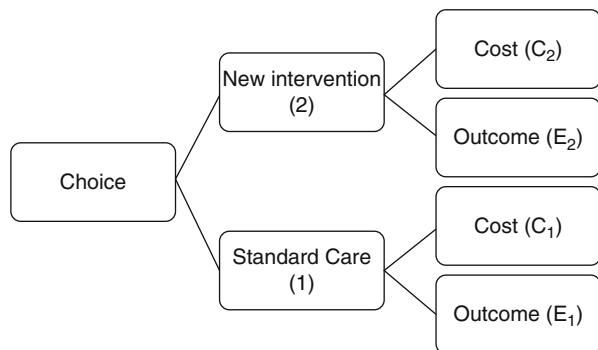
or improve prior to their death in 4 years). For this individual their  $\text{QALYs} = 4 \times 0.5 = 2$ . However, they may have an operation for their health condition, which will improve their health state giving a utility value of 0.9 and life expectancy of 8 years (again at a constant health state). In this case their  $\text{QALYs} = 8 \times 0.9 = 7.2$ . Thus the QALY gain from the operation  $= 7.2 - 2 = 5.2$  QALYS. This is shown in Fig. 1.2.

The benefit of the use of CUA rather than CEA using the natural units described in Sect. 1.3.2 is that it permits comparisons between as well as within health care programmes. If we are comparing interventions using, for example, cancers detected, we can only compare interventions designed to detect cancers. Using QALYs we can make a comparison between, for example, an intervention designed to detect cancer and an intervention designed to lower blood pressure. The analysis also has the advantage of incorporating quality of life and takes account of preferences for different health states. However, the analysis is limited to health benefits (that can be measured by the outcome), and there are challenges in deriving health state utilities (Brazier et al. 2007; Nord 1999; Dolan 2000; Devlin et al. 2012).

## 1.4 Incremental Cost Effectiveness Ratios (ICERs)

In both CEA and CUA, comparisons are made between two or more interventions. Figure 1.3 shows the typical structure of the analysis when two interventions are compared.

**Fig. 1.3** Structure of cost effectiveness analysis comparing two interventions



**Fig. 1.4** Example of a cost effectiveness plane

In order to accurately reflect the opportunity cost of the new intervention a comparison should be made against the next best alternative (the control or comparator). This is typically, but not always, current standard care. The incremental costs and incremental effectiveness of an intervention can be graphed on a cost effectiveness plane. On this diagram, each point represents how different the intervention is to the comparator.

The cost effectiveness plane consists of four quadrants in which the incremental costs are on the vertical axis and the incremental effect on the horizontal axis (Fig. 1.4). Any point in the North East quadrant has higher costs and is more effective; points in the South East have lower costs and are more effective; in the North West higher costs and less effective; and in the South West lower costs but less effective.

The outcome of the analysis is often presented in shorthand as an incremental cost effectiveness ratio (ICER). The ICER measures the incremental cost of an activity

relative to incremental next best alternative, divided by the incremental effectiveness between those same two alternatives. The ICER calculation is given by the formula

$$\text{ICER} = \frac{C_2 - C_1}{E_2 - E_1} = \frac{\Delta C}{\Delta E}$$

where:

$C_2$  is the cost under the intervention of interest.

$E_2$  is the effectiveness under the intervention of interest.

$C_1$  is the cost under the comparator.

$E_1$  is the effectiveness under the comparator.

On the cost effectiveness plane, any intervention is identified by its incremental cost and incremental effectiveness against the comparator or control. When compared to itself, the comparator/control has a zero incremental cost and a zero incremental effectiveness; on the diagram it can be represented by the origin or point where the horizontal and vertical axes intersect. When we compare the intervention and control, we can draw a line between the origin and the intervention point. The ICER is the gradient of this line.

If the ICER lies in the South East quadrant, the ICER is negative and the new intervention is cost effective; it is said to dominate the control or comparator (more effective and less costly than the comparator or control). If the ICER lies in the North West quadrant, the ICER is negative, but the new intervention will not be cost effective; it is said to be dominated (more costly and less effective than the comparator or control). For ICERs that are in the remaining two quadrants (the South West and North East), the ICER is positive and there are trade-offs to consider. In the North East, whilst the costs are higher, the intervention is more effective; similarly in the South West, whilst the costs are lower, the intervention is not as effective as the control or comparator.

So what happens when the ICER falls into the North East or South West quadrants? In these cases, the ICER is compared to a ceiling ratio, which is a level of the ICER which an intervention must meet to be regarded as cost effective. Sometimes, this ceiling ratio is taken to represent a willingness to pay for each outcome, for example, as a willingness to pay per QALY. In this case, the ceiling ratio represents the value placed on each unit of outcome by society or by a decision maker. More often, however, the ceiling ratio is known as the cost effectiveness threshold, and in this case, the ceiling ratio represents the opportunity cost of spending money on other health treatments. Although the interpretation of the ceiling ratio (denoted by the Greek letter lambda,  $\lambda$ ) is very different in each case, a common decision rule applies to both.

If the new intervention is more effective but at a higher cost, then:

$\text{ICER} < \lambda$ , the activity is cost effective.

$\text{ICER} > \lambda$ , the activity is not cost effective.

If the new intervention is less costly but less effective, then:

$\text{ICER} > \lambda$ , the activity is cost effective.

$\text{ICER} < \lambda$ , the activity is not cost effective.

In the case where we have an ICER equal to the ceiling ratio, the productive efficiency of the health system will be the same regardless of whether or not the intervention is chosen. In this (highly unlikely) case, the standard decision rule does not give any guidance. Chapter 4 (Sect. 4.4) discusses the cost effectiveness plane further in relation to problems with ICERs.

### ***1.4.1 Simple and Extended Dominance***

When we are comparing two alternative interventions or courses of action as we have done above, it is relatively simple to determine that if the new intervention or activity is less costly and more effective (in the South East quadrant), it is dominant (or dominates). This is known as simple dominance. But what happens if we have more than two alternatives?

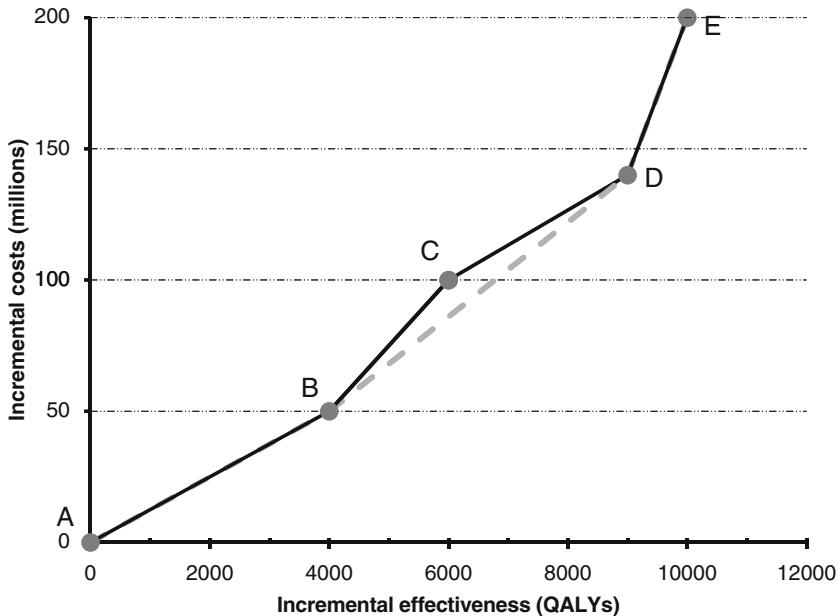
With more than two alternatives, the concept of simple dominance will still apply; one or more may be less costly and more effective than others. In addition we can determine whether a combination of two or more options is less costly and more effective. For this case, we will consider an example with five diagnostic tests, A–E, that could be provided to 500 people. These options are presented in Table 1.1 in order of increasing effectiveness. For ease of interpretation, we assume that the least costly, least effective Option A represents our ‘standard care’ and so is used as the comparator. Figure 1.5 identifies these options in terms of the incremental costs and effects. As the comparator, Option A appears at the intersection of the incremental cost and incremental effects axis with Options B–E representing progressively more effective (although not necessarily more cost effective) options. Option B is compared to A, and then C is compared to B and so on to calculate the ICERs reported in the final column of Table 1.1.

None of the tests can be eliminated due to simple dominance as none are less costly and more effective than any of the others. In addition we can determine whether a combination of two or more options might be less costly and more effective than one of our existing options.

On Fig. 1.5, we have already identified that the slope of a line drawn between Options A and B represents the ICER between them. There is another interpretation available for this line, since this also represents the lines that are possible by considering a combination of Options A and B, where we allocate a fixed proportion of people (at random) to these two options. Here, the point at ‘Option A’ represents

**Table 1.1** ICERs between five diagnostic tests (Options A–E)

Test	Cost (\$M)	Effect (QALYs)	ICER ( $\Delta C/\Delta E$ )
A (standard care)	50	10,000	
B	100	14,000	\$12,500 per QALY
C	150	16,000	\$25,000 per QALY
D	190	19,000	\$13,333 per QALY
E	250	20,000	\$60,000 per QALY



**Fig. 1.5** Incremental cost effectiveness ratios for diagnostic tests compared to standard care

**Table 1.2** ICERs between four diagnostic tests (Option C removed)

Test	Cost (\$M)	QALYs	ICER ( $\Delta C/\Delta E$ )
A (standard care)	50	10,000	
B	100	14,000	\$12,500 per QALY
D	190	19,000	\$18,000 per QALY
E	250	20,000	\$60,000 per QALY

where 500 people are given Option A and 0 are allocated to Option B; Option B represents the case where these numbers are reversed. If 250 people were allocated to Option A and 250 to Option B, we would expect outcomes that lie halfway between Options A and B.

In a similar way, we can also consider combinations between any two (or more) treatments. For example, suppose we allocated 250 people to Option B and 250 to Option D; in this case we would expect an outcome that is the average of B and D. This option would have expected costs of \$145M and effectiveness of 16,500 QALYs. This compares to a cost of \$150M but only 16,000 QALYs if we choose Option C. In this case, a combination (and in fact many potential combinations) between Options B and D would be expected to dominate Option C, and so we say that C is extended dominated. When we consider the ICER between Options B and D, by implication we also consider all the combinations between them and in so doing already consider outcomes that are better than Option C. For this reason, we have enough reason to ignore Option C. Table 1.2 provides an updated list of

alternatives where we have recalculated the incremental cost and effect of each remaining option relative to all the previous options.

In this case, the ICERs still range in value from \$12,500 to \$60,000 per QALY, although this is not always going to be the case after removing extended dominated alternatives. The decision maker can now use the results to decide which test or tests to provide. As highlighted earlier, the decision rule is based on the cost effectiveness threshold or the willingness-to-pay threshold ( $\lambda$ ). For example, if the cost effectiveness threshold is \$20,000 per QALY, this would imply that E is unlikely to be provided; whilst Tests B and D are both below the threshold, D is more cost effective than B. We pick the test with the highest ICER that is below the threshold.

If we had ten alternatives (Options A–J) ordered by effectiveness, then calculating the ICERs would require that we consider nine pairs of options: Options A to B, B to C, and so on until I to J. If we had to compare all possible alternative combinations, then this would be a lot of work. Instead of doing this it is enough just to look at the ICER column. If we do not have any dominated or extended dominated options and we consider progressively more effective treatments, we will find that the ICERs increase with every comparison. In Table 1.2, for example, we have a jump from \$12,500 to \$18,000 to \$60,000 per QALY. As we highlighted previously, lower ICERs are not necessarily more cost effective.

Where we have dominated or extended dominated options, we will not find ICERs that rise with increasing effectiveness. For example, either we might find that there are some negative ICERs (so that there are some items that should be removed by simple dominance) or we might find that the ICERs increase and then decrease. In the case of Table 1.1, we found that the ICER increased when we moved from B to C and then decreased when we moved from C to D; this demonstrates that C was extended dominated even before we checked C with a combination of B and D. Using this rule of thumb, it is possible to identify dominated or extended dominated options very quickly. It is important to say at this point that after each dominant or extended dominant strategy is removed; all ICERS must be recalculated and checked again for dominance or extended dominance strategies. You should repeat this until ICERS consistently rise with increased effectiveness.

#### 1.4.2 *The Net Benefit Approach*

It is important to measure an ICER correctly to compare the new intervention or intervention of interest with all relevant alternatives. Use of an inappropriate comparator can lead to bias or give misleading results, and there are mathematical difficulties associated with use of a ratio. An alternative to using the ICER is the net benefit approach (Stinnett and Mullahy 1998), in which either costs are transformed to be in the same units as effectiveness or effectiveness is transformed into the same units as costs. Suppose, as in the examples of this chapter, that our effectiveness unit is given by the QALY and our costs by dollars. In order to transform our measures

of effectiveness (in QALYs) to cost units (\$), we need to have an ‘exchange rate’ between the two: this \$/QALY figure is the ceiling ratio.

If we multiply our measure of effectiveness by the ceiling ratio, we convert the effectiveness to a monetary unit and can define the Net Monetary Benefit (NMB) as the difference between monetised incremental effectiveness ( $\lambda \times \Delta E_i$ ) and monetary incremental costs ( $\Delta C_i$ ) in terms of the currency unit (e.g. \$). Similarly, if we work in the opposite direction, we can convert costs into effectiveness terms by dividing through by the ceiling ratio. In this case, the Net Health Benefit (NHB) is given by the difference between incremental effectiveness ( $\Delta E_i$ ) and health equivalent of the costs ( $C/\lambda$ ), where this will be measured in terms of the effectiveness unit (e.g. QALYs). Whether we are interested in NMB or NHB, the most cost effective treatment is the one with the highest Net Benefit. We discuss the NB framework in more detail in Chap. 11.

One of the difficulties with the net benefit approach is identifying a value of  $\lambda$  to be used. This might occur where there is a lack of previous decisions using this unit, especially where the CEA uses an effectiveness unit that is quite specific to a particular clinical area (e.g. if we wished to find a willingness to pay per filled, missing or decayed deciduous tooth) or where they may not choose to identify a credible ceiling ratio (CADTH 2006). In such cases, an alternative to presenting ICERs that has been gaining popularity over the past few years is the cost effectiveness acceptability curve (CEAC) (van Hout et al 1994). CEACs account for statistical uncertainty in the ICER and uncertainty in the ceiling ratio. They show the probability that an intervention is cost effective for different levels of the ceiling ratio ( $\lambda$ ) but crucially do not require the analyst to know which value of the ceiling ratio the decision maker might wish to use. There is an extensive literature describing the challenges involved in characterising the uncertainty around estimated ICERs (Polsky et al. 1997; Campbell and Torgerson 1999). Uncertainty is discussed in more detail in Chap. 4.

## 1.5 Summary

- Economics is based on the premise of scarcity; any spending choices or decisions that are made about health care provision incur an opportunity cost.
- Economic evaluation is a form of comparative analysis; it allows interventions to be assessed in terms of their benefits and costs to provide information to allocate resources efficiently.
- CEA uses one measure of effectiveness, for example, cancers detected. This restricts comparison across analyses with different outcomes.
- Within CUA (often also referred to as CEA) the effect is measured in terms of healthy years combining life years gained with the quality of those life years. Measures include QALYs which are widely accepted as the reference standard (Gold et al. 1996; NICE 2013).
- When carrying out a CEA or CUA, both the incremental cost and incremental effectiveness are important. The outcome of the analysis is often presented as an ICER and is calculated:  $ICER = (C_2 - C_1)/(E_2 - E_1) = \Delta C/\Delta E$ .

- ICERs may be shown graphically on a cost effectiveness plane. When comparing the intervention and control, on the cost effectiveness plane, we can draw a line between the origin and the intervention point. The ICER is the gradient of this line. If the ICER lies in the South East quadrant, the new intervention is cost effective and dominates; in the North West quadrant, the new intervention will not be cost effective; it is dominated. For ICERs in the remaining two quadrants, there are trade-offs to consider.
- The net benefit approach is an alternative to using the ICER in which either costs are transformed to be in the same units as effectiveness or effectiveness is transformed in order to be in the same units as costs.

## References

- Brazier JE, Ratcliffe J, Salomon J, Tsuchiya A (2007) Measuring and valuing health benefits for economic evaluation. OUP, Oxford
- Briggs AH, O'Brien BJ (2001) The death of cost-minimization analysis? *Health Econ* 10(2):179–184
- Campbell MK, Torgerson D (1999) Bootstrapping: estimating confidence intervals for cost effectiveness ratios. *QJM* 92(3):177–182
- Canadian Agency for Drugs and Technologies in Health (2006) Guidelines for the economic evaluation of health technologies, 3rd edn. Canadian Agency for Drugs and Technologies in Health, Ottawa
- Dakin R, Wordsworth S (2013) Cost minimisation analysis revisited. *Health Econ* 22(1):22–34
- Department of Health (2010) A new value-based approach to the pricing of branded medicines – a consultation. Department of Health, London
- Devlin N, Buckingham K, Shah KK, Tsuchiya A, Tiling C, Wilkinson G, van Hout B (2012) A comparison of alternative variants of the lead and lag time TTO. *Health Econ* 22(5):517–532
- Dolan P (2000) The measurement of health-related quality of life for use in resource allocation decisions in health care. *Handb Health Econ* 1(B):1723–1760
- Drummond MF, Stoddart GL, Torrance GW (1987) Methods for the economic evaluation of health care programmes, 1st edn. Oxford Medical Publications, Oxford
- Gold MR, Siegel JE, Russell LB, Weinstein MC (1996) Cost effectiveness in medicine and health. Oxford University Press, Oxford
- Johansson PO (1991) An introduction to modern welfare economics. Cambridge University Press, Cambridge
- National Institute for Health and Care Excellence (2013) Guide to the methods of technology appraisal. National Institute for Health and Care Excellence, London
- Nord E (1999) Cost-value analysis in health care: making sense out of QALYs. Cambridge University Press, Cambridge
- Polsky D, Glick H, Willke R, Schulman K (1997) Confidence intervals for cost effectiveness ratios. A comparison of four methods. *Health Econ* 6:243–252
- Stinnett A, Mullahy J (1998) Net Health Benefits: a new framework for the analysis of uncertainty in cost effectiveness analysis. *Med Decis Making* 18(2 Suppl):S68–S80
- Sugden R, Wiliams A (1978) The principles and practice of cost benefit analysis. Oxford University Press, Oxford
- Van Hout BA, Al M, Gordon GS, Rutten FFH (1994) Costs, effects and C/E-ratios alongside a clinical trial. *Health Econ* 3:309–319
- World Bank (1993) World Development Report 1993: Investing in Health. New York: Oxford University Press. © World Bank. <https://openknowledge.worldbank.org/handle/10986/5976>. License: CC BY 3.0 IGO

# **Chapter 2**

## **Finding the Evidence for Decision Analytic Cost Effectiveness Models**

**Abstract** Decision analytic cost effectiveness models are a mechanism for synthesising disparate evidence on the safety, effectiveness and cost of alternative health technologies, along with evidence on the epidemiology of the disease, the technologies designed to treat and the processes that the health care system uses to manage people over the disease course. The methods for identifying the evidence that the model will synthesise are an important determinant of the quality of the results of a model. This chapter will familiarise you with the good practice in the processes of identifying evidence for cost effectiveness models and evidence sources, factors to consider when selecting sources, and it will show you how to construct a search strategy.

### **2.1 Introduction**

In Chap. 1, we provided an overview of the different types of economic evaluation with particular focus on cost effectiveness analysis (CEA) (including cost-utility analysis (CUA)). CEA is typically undertaken alongside clinical trials, and there are clear advantages in doing so, specifically that the trial provides the opportunity to observe and collect resource use and outcomes for a given patient population in the same conditions in a single study which may not be available at any other time (Johnston et al. 1999). However, use of single trial data alone limits analyses to the population within the trial, the interventions assessed and the time over which the trial participants are followed up (NICE 2013). Using decision analytical modelling within cost effectiveness analyses allows us to break out of these confines and to tailor our analyses to the question we wish to answer. The modelling process facilitates use of multiple sources of data to answer questions relating to resource allocation that are typically beyond the scope of a single clinical trial.

The methods for identifying the evidence that the model will synthesise are an important determinant of the quality of the results of a model. They are also an area of independent expertise and research. The parent discipline for the

professionals in this area is most likely health librarianship, but increasingly the leading practitioners in this field are known as health information specialists. Whatever their title, their expert input is essential to any team engaged in serious decision analytic cost effectiveness modelling.

Detailed training in information searching is beyond the scope of this chapter, but it is important that when undertaking cost effectiveness modelling, you have sufficient understanding of the key methods and sources employed to be able to engage effectively with information specialists. Therefore, before getting into the details of how to build decision analytic cost effectiveness models, we use this chapter to familiarise you with the good practice in the processes of identifying evidence for cost effectiveness models.

These processes and resources differ to those used in identifying evidence for reviews. Many reviews start with establishing the review question followed by a literature search aimed at identifying studies directly relevant to a single question. Cost effectiveness models evolve as the care pathways and parameters are identified, and hence it is difficult to define the search questions at the start of the model. A single literature search at the start of the economic model is unlikely to capture all the data required (Kaltenthaler et al. 2014). Model searches can be less exhaustive and more targeted than review searches since their purpose is to identify adequate evidence to populate the model, rather than retrieve all studies (Glanville and Paisley 2010). This chapter focusses on the requirements and methods for the key searches often undertaken by the information specialist. Bear in mind further individual searches will be undertaken to help populate the model when the clinical pathways and model parameters are identified. The modelling process uses a wide range and variety of resources, which means lots of targeted searching! Analysts and information specialists should both be aware of the iterative search processes that need to be flexible to respond to the information requirements of the model. Whilst the information specialist is pivotal in formulating and running the searches to locate information, this by no means precludes searching by the analyst; in fact models typically use evidence from searches by the analyst and the information specialist as well as information provided by experts and key contacts.

The chapter is structured as follows: Section 2.2 outlines evidence sources and factors to consider when selecting sources. Section 2.3 describes how a search strategy is constructed from a research question. These methods are applicable for designing searches in the subsequent sections. Section 2.4 describes searching for existing cost effectiveness models. Section 2.5 is focussed on identifying evidence on the epidemiology of a disease and the safety and effectiveness of technologies. Section 2.6 considers how analysts should search for evidence on health-related quality of life and health state preferences, and Section 2.7 considers how to find evidence on the resource use and costs of care. In each section we identify pertinent sources of evidence to search and describe validated search strategies where they are available. Section 2.8 describes methods to log and report search activity. In Section 2.9 we briefly review the quality assessment tools that are available for

different types of evidence. The key points covered in the chapter are summarised in Section 2.10.

## 2.2 Choosing Resources to Search for Evidence

Evidence used in models of cost effectiveness comes from many sources including evidence syntheses, expert judgement, observational studies, randomised controlled trials (RCTs), other clinical studies, reference sources and routine data sources (Paisley 2010). International, national and local databases and websites provide access to such research literature and data. The sections in this chapter list key resources, but updated links to wider sets of resources are available on the Health Economics Core Library Recommendations website (AcademyHealth 2011) and the Health Economics Resources website (University of York Library 2015).

Useful research studies may be found in multidisciplinary resources (e.g. Scopus), subject-specific resources (e.g. CINAHL for nursing literature) and evidence-based resources (e.g. McMaster Health Systems Evidence) (McMaster University 2015). The choice of databases to search depends on the database accessibility, subject coverage, currency and type of data/reports covered. The facilities to download records in bulk, availability of full-text papers and the researcher's search expertise can also influence this decision. Some databases are freely available to the public, and others will be accessible via institution database subscriptions. PubMed is the freely available version of MEDLINE. Many institutions subscribe to MEDLINE via database hosts such as Ovid or EBSCO. PubMed has a slightly larger coverage than the subscription version of MEDLINE but has fewer advanced search features and download options (National Library of Medicine 2014). For many 'health economics' searches, the three key databases to search have been MEDLINE, Embase and NHS Economic Evaluations Database (NHSEED). Royle and Waugh (2003) demonstrated that searching all three identified 94.8 % of the studies used in cost effectiveness sections of 20 Technology Assessment Reports. However, from 2015 NHSEED is no longer being updated. MEDLINE and Embase remain the two key databases but multi-disciplinary databases e.g. Web of Science Core Collection may be worth searching for records of cost studies that are not available in MEDLINE and Embase.

Databases and websites containing population and health care use statistics, unit costs and health care guidelines also provide valuable data for cost effectiveness models. Some data repositories provide data free of charge, whilst others charge.

Librarians and information specialists are well placed to advise on the most suitable databases for the research question, their subscriptions and their ease of use and can offer support in developing search skills.

## 2.3 Designing Search Strategies

There are standard approaches to developing a literature search in order to gather evidence relevant to a research question. The guidance here briefly outlines that approach and applies it to identifying relevant evidence for cost effectiveness models, primarily from bibliographic databases such as MEDLINE. Further, more definitive guidance on the process of developing a search strategy can be sought from your institution's library service and core texts (Aveyard 2010; Booth et al. 2011).

Developing a literature search strategy requires identifying words, phrases and index terms (database subject headings) which are present in studies relevant to the research question. Searches (or queries) are then run on the database for each term. The results of the searches (retrieved records) are then combined using 'Boolean' logic to produce a final set of records of references that are relevant to the question. An initial plan of the search helps identify possible terms and phrases to use and the logic combinations of searches. It is useful to break the question down into distinct search concepts for which terms are collected and combined. For example, the question 'What are the existing economic evaluations of hearing aids for elderly people with hearing impairments?' contains the search concepts (A) 'hearing aid', (B) 'elderly people', (C) 'economic evaluations' and (D) hearing impairment. It is advisable to select the two or three concepts that are most likely to be successful in retrieving relevant papers. In this example the concepts A, B and C should retrieve economic evaluations of elderly people using hearing aids. The 'hearing impairment' concept is the most problematic to search as there are many terms and phrases that can indicate hearing impairments. It could be time-consuming to collect and enter all hearing impairment terms into the database. Also the search for 'hearing aids' is only likely to find studies of people with hearing impairment, making it an unnecessary search concept to use. Noting possible terms and phrases for each search concept is useful to collate search ideas whilst maintaining the search concept structure of the search. Table 2.1 is an example table of search concepts and terms.

Identifying concepts within a research question is the same approach as using PICO elements to plan searches for a systematic review literature search

**Table 2.1** Search concepts and terms for the question 'What are the existing economic evaluations of hearing aids for the elderly?'

Search concept	(A) Hearing aid	(B) Elderly people	(C) Economic evaluations
Search terms	Hearing aid	Elderly	Economic evaluations
	Hearing device	Old age	Cost effectiveness analysis
	Deaf aid	Elderly	Cost-benefit analysis
	Ear aid	Pensioner	Cost-utility analysis
		Retired	
		Over 65	
		Older person	
		Older people	

(Lefebvre et al. 2011). The PICO tool outlines the four elements of a well designed effectiveness question. P is the patient problem or population, I the intervention, C the comparison and O the outcome (Sackett et al. 1997). In the example search here, P is elderly people with hearing impairment, I is hearing aid, C is unknown and O is improved hearing. We can include a further study type concept of ‘economic evaluations’.

Some search concepts may have search filters (ready-made searches) available for use. These can be applied to the search instead of developing a particular search concept. Search filters are usually designed to retrieve types of study, for example, RCTs or observational studies. They may be validated and some are published in peer-reviewed papers (InterTASC ISSG 2015a). The Information Specialists’ Sub-Group (ISSG) Search Filters Resource provides more detailed information about search filters and links to filters for a wide range of study types. Section 2.5.2 provides example searches that include the RCT search filter along with an explanation of how to incorporate it into a search.

Once each search concept has been searched in the database, the final combination search retrieves references most likely to be relevant to your research question. The final search lines for each search concept can be seen in Fig. 2.1. Line 6 combines all the hearing aid terms together, line 12 combines all the elderly terms together and line 18 combines all the economic evaluation terms together. The final combination in line 19 instructs the database to retrieve only those references that have a ‘hearing aid’ term, an ‘elderly’ term and an ‘economic evaluation’ term within the same reference.

Some databases have options to limit a search to a study or publication type. These limits are rarely 100 % accurate but can quickly limit the search to sufficient evidence for the model. PubMed and Ovid databases have an ‘additional limits’ feature containing search limits including publication types, ages (in respect of the population of interest), languages and clinical queries. Each of the following search sections will include details of any relevant limit options available on key databases. Limits to identify clinical effectiveness studies are more common than other study types, and they are described in Sect. 2.5.3.

Each chapter section will outline an example search question, the search concepts and suggested terms.

## 2.4 Searching for Existing Cost Effectiveness Models

A targeted literature search to identify existing cost effectiveness models at the start of the modelling process can help avoid duplicating work previously undertaken. Reviewing existing cost effectiveness models may be useful to inform the development of your model. Such studies may describe clinical pathways, possible parameters and potential data sources. However, models should be scrutinised and carefully considered before any data or model design is incorporated into the new model. The existing models may be addressing different questions,

Ovid MEDLINE(R) <1946 to September Week 3 2014>

Search Strategy:

- 
- 1 Hearing Aids/ (6859)
  - 2 hearing aid\*.ti,ab. (6262)
  - 3 hearing device\*.ti,ab. (282)
  - 4 ear aid\*.ti,ab. (45)
  - 5 deaf aid\*.ti,ab. (2)
  - 6 or/1–5 (8562) [hearing aids]
  
  - 7 exp Aged/ (2396918)
  - 8 (aged or elderly or geriatric\*).ti,ab. (507035)
  - 9 ("old\* person" or "old\* people" or "old age\*").ti,ab. (35665)
  - 10 (pensioner\* or retired).ti,ab. (4705)
  - 11 ("over 65 y\*" or "over 70 y\*" or "over 80 y\*").ti,ab. (6587)
  - 12 or/7–11 [Aged] (2653674)
  
  - 13 exp "Costs and Cost Analysis"/ (185208)
  - 14 economic evaluation\*.ti,ab. (6249)
  - 15 Cost effectiveness.ti,ab. (33326)
  - 16 Cost-benefit.ti,ab. (6906)
  - 17 Cost-utility.ti,ab. (2337)
  - 18 or/13–17 [Economic evaluations] (200092)
  
  - 19 6 and 12 and 18 (62)

#### **Key**

- Exp    explode subject heading (searches additional, closely related but more specific terms)
- /      Medical Subject Heading (MeSH)
- \*
- " ... "    searches for the two or more words within the quotation marks as a phrase
- .ti, ab    searches title and abstract
- Or/1–3    set combination 1 or 2 or 3

Each chapter section will outline an example search question, the search concepts and suggested terms.

**Fig. 2.1** Search strategy for the question ‘What are the existing economic evaluations of hearing aids for the elderly?’

and their data sources may need to be verified as appropriate to the new model (Kaltenthaler et al. 2014).

#### **2.4.1 Where to Look**

Studies describing cost effectiveness models can be found in health journals, HTA reports and conference abstracts. It is important to note that conference abstracts may lack the detail required by the analyst; however, there may be a subsequent peer-reviewed paper that contains further details, or you can contact the authors of

**Table 2.2** Search concepts and terms for the question 'What are the existing cost effectiveness models of Herceptin use in breast cancer?'

Search concept	(A) Breast cancer		(B) Herceptin	(C) Cost effectiveness model
Search terms	Breast	Cancer*	Trastuzumab	Economic model*
	Mammary	Neoplasm*	Herceptin	Cost effectiveness model*
		Tumor*	Herclon	Markov model*
		Carcinoma*		Decision model*
		Adenocarcinoma*		Discrete event model*
		Malignan*		Discrete event simulation*
		Sarcoma*		Patient level simulation*
				Microsimulation*

\*Denotes truncation

the conference abstract to request further information. Searching a combination of MEDLINE and Embase may identify a good proportion of existing cost effectiveness model studies. Further studies could be sought from specialist economic databases, for example, EconLit, and RePEc. Databases covering more specific health subjects may also be useful depending on the research question. For example, if developing a model for schizophrenia, it would be advisable to search the mental health database PsycINFO for existing cost effectiveness models of schizophrenia care pathways.

#### 2.4.2 *Search Strategy, Concepts, Terms and Combinations*

The search strategy for a search for existing cost effectiveness models is likely to include search concepts for the 'Patient/Population' and 'Intervention' (as outlined in PICO in Sect. 2.3) under consideration by the model. A further search concept would limit studies to just those mentioning cost effectiveness models. For example, the question 'What are the existing cost effectiveness models of Herceptin use in breast cancer?' contains the search concepts (A) 'breast cancer', (B) 'Herceptin' and (C) 'cost effectiveness model'. Table 2.2 illustrates the three search concepts plus terms identified as potentially useful for the search.

Individual searches for each search concept are developed by searching for words and phrases in the title and abstracts of the database record and index terms (e.g. MeSH in MEDLINE, EMTREE in Embase). The individual searches are then combined together to identify studies that have a term from each of the concepts. Figure 2.2 illustrates the structure of an example targeted search for cost effectiveness models of Herceptin in breast cancer. The final search line combines the three search concepts using the 'AND' Boolean combination operator, i.e. A AND B AND

Ovid MEDLINE(R) <1946 to September Week 4 2014>Search Strategy:

- 
- 1 exp breast neoplasms/ (223494)
  - 2 (breast adj5 (neoplasm\* or cancer\* or tumour\* or tumor\* or carcinoma\* or adenocarcinoma\* or malignan\* or sarcoma\*)).ti,ab.(215009)
  - 3 (mammary adj5 (neoplasm\* or cancer\* or tumour\* or tumor\* or carcinoma\* or adenocarcinoma\* or malignan\* or sarcoma\*)).ti,ab.(27024)
  - 4 or/1–3 [Breast Cancer] (282567)
  - 5 trastuzumab.ti,ab.(4601)
  - 6 herceptin.ti,ab.(1384)
  - 7 herclon.ti,ab.(0)
  - 8 or/5–7 [Herceptin] (5246)
  - 9 exp models, economic/ (10465)
  - 10 markov chain/ (10270)
  - 11 decision support techniques/ (12863)
  - 12 (econom\* adj2 model\*).ti,ab.(2650)
  - 13 ("cost effectiveness" adj2 model\*).ti,ab.(969)
  - 14 (markov\* adj5 model\*).ti,ab.(7743)
  - 15 (decision\* adj8 model\*).ti,ab.(10953)
  - 16 (discrete event\* adj8 model\*).ti,ab.(308)
  - 17 (discrete event\* adj5 simulat\*).ti,ab.(381)
  - 18 (patient level adj8 simulat\*).ti,ab.(38)
  - 19 microsimulat\*.ti,ab.(429)
  - 20 or/9–19 [Cost effectiveness Models] (43702)
  - 21 4 and 8 and 20 (44)

#### Key

- Exp    explode subject heading (searches additional, closely related but more specific terms)
- /       Medical Subject Heading (MeSH)
- Adj 5   terms must be within 5 words of each other
- \*truncated term identifying terms with the same stem
- " ..."   searches for the two or more words within the quotation marks as a phrase
- .ti, ab   searches title and abstract
- Or/1–3   set combination 1 or 2 or 3

**Fig. 2.2** Example search for cost effectiveness models of Herceptin in breast cancer (in Ovid MEDLINE®)

C. This ensures that the final set of results contains a breast cancer term (A) AND a Herceptin term (B) AND a cost effectiveness model term (C) within the database record.

#### 2.4.3 *Search Filters, Database Limits and Clinical Queries*

A validated ‘cost effectiveness model’ search filter is not currently available. There are no ‘limits’ or clinical queries available in databases to limit a search to studies of cost effectiveness models.

## 2.5 Searching for Clinical Evidence

### 2.5.1 *Finding the Evidence on Incidence, Prevalence and Natural History of a Disease*

Information is required to accurately predict the incidence and describe the prevalence of a disease in a given population within the model. Evidence on the natural history of a disease provides a view over time of what resources may be called upon and can indicate alternative care pathways.

Large sets of ‘routine data’ can provide a current and real-world evidence base of the disease’s epidemiology. Table 2.3 lists some key national and international health statistics resources, for example, in England, the Health and Social Care Information Centre (HSCIC) provides access to UK Hospital Episode Statistics and the Compendium of Population Health Indicators. Charities, pressure groups and research centres can also provide quick access to epidemiology data.

Research into the epidemiology of a disease can also be sought from books, journal papers and the databases they are indexed in. Longitudinal and observational studies are a good source of epidemiology data. There is evidence to suggest that searching two databases (MEDLINE and Embase) could be adequate for the identifying epidemiology data for the a model (Royle et al. 2005). Searches of appropriate subject-specific databases, e.g. PEDro for physiotherapy (The George Institute for Global Health 2014), or non-English language databases, e.g. LILACs (BIREME 2015), may retrieve studies relevant to the particular research question that have not been found elsewhere.

Brief, targeted literature searches for epidemiology studies consist of a search concept for the disease or health care condition (e.g. hearing loss), and a second

**Table 2.3** Selected health statistics resources

Resource	Geographic coverage	URL (accessed 29-9-14)
Health Data Tools and Statistics	International and the USA	<a href="http://phpartners.org/health_stats.html">http://phpartners.org/health_stats.html</a>
OECD Statistics – Health	International	<a href="http://stats.oecd.org/">http://stats.oecd.org/</a>
Australian Institute of Health and Welfare	Australia	<a href="http://www.aihw.gov.au/">http://www.aihw.gov.au/</a>
Statistics Canada	Canada	<a href="http://www5.statcan.gc.ca/subject-sujet/theme-theme.action?pid=2966&amp;lang=eng&amp;more=0&amp;MM">http://www5.statcan.gc.ca/subject-sujet/theme-theme.action?pid=2966&amp;lang=eng&amp;more=0&amp;MM</a>
EC health indicators	European regions	<a href="http://ec.europa.eu/health/indicators/echi/index_en.htm">http://ec.europa.eu/health/indicators/echi/index_en.htm</a>
Health and Social Care Information Centre (HSCIC)	UK	<a href="http://www.hscic.gov.uk/">http://www.hscic.gov.uk/</a>
National Center for Health Statistics	USA	<a href="http://www.cdc.gov/nchs/">http://www.cdc.gov/nchs/</a>

**Table 2.4** Search concepts and search terms for the question 'What is the incidence and prevalence of hearing loss in the UK?'

Search concept	(A) Epidemiology	(B) Hearing loss	(C) UK
Search terms	Epidemiology	Hearing loss	Great Britain
	Incidence		UK
	Prevalence		England
	Natural history		Ireland
	Risk factors		Scotland
			Wales
			National Health Service

concept is epidemiology (to cover incidence, prevalence and natural history). A third concept could be a geographic region or another contextual concept. Table 2.4 shows search concepts and terms that could be used to build a search strategy for identifying studies of natural history, incidence and prevalence of hearing loss within the UK. This strategy uses a limited number of focussed MeSH headings and searches for terms within the title (not abstract) to limit the results to studies that are most likely to be relevant. The search is less exhaustive than the cost effectiveness model search because it only needs to find a small number of highly relevant papers that contain the required epidemiology data.

Figure 2.3 illustrates the search strategy. Each search concept is searched before a final combination of A (line 7) and B (line 9) and C (line 17) is made. When searching MEDLINE, the epidemiology subheading can be used with the health care condition MeSH. In this example search line 10 (*exp \*Hearing Loss/ep*) retrieves studies indexed with the 'Hearing Loss – Epidemiology' subject heading. This search line is a combination in itself of A (epidemiology) and B (hearing loss).

There are no validated search filters for identifying epidemiology studies at present. Published research in searching for epidemiology studies mentions using the sub-heading 'Epidemiology' when searching the health care condition subject heading (line 10 in Fig. 2.3). Also, specific epidemiology subject headings can be used (see lines 1–5 in Fig. 2.3). The 'causation-aetiology' clinical query can be added as an additional limit in MEDLINE, PubMed and Embase to identify risk studies.

### 2.5.2 *Finding the Evidence on the Clinical Effectiveness of Health Interventions*

Clinical effectiveness evidence provides the clinical outcome data required by the model and may also include data on safety, adverse events and complications following the health intervention. Results of RCTs, evidence syntheses and systematic reviews of RCTs are considered to offer the highest quality evidence of clinical effectiveness of health interventions. A literature search for syntheses, reviews and RCTs may identify sufficient evidence to inform the clinical effectiveness data

## Ovid MEDLINE(R) &lt;1946 to September Week 3 2014&gt;Search Strategy:

- 
- 1    \*Epidemiology/ (9324)
  - 2    \*incidence/ (415)
  - 3    \*prevalence/ (555)
  - 4    \*risk factors/ (935)
  - 5    \*Natural History/ (426)
  - 6    (epidemiology or incidence or prevalence or "risk factor\*" or "natural histor\*").ti. (265815)
  - 7    or/1–6 (273588)
  - 8    exp \*Hearing Loss/ (42381)
  - 9    7 and 8 (673)
  - 10    exp \*Hearing Loss/ep[Epidemiology](1561)
  - 11    9 or 10 [Hearing Loss Epidemiology] (1922)
  - 12    exp great britain/ (308700)
  - 13    ("united kingdom\*" or uk or "U.K." or "UK." or "U.K" or britain).ti. (30789)
  - 14    (british or english or scottish or welsh or irish).ti. (24557)
  - 15    (england or wales or scotland or ireland).ti. (23640)
  - 16    (nhs or "national health service").ti. (8234)
  - 17    or/12–16 [UK] (334141)
  - 18    11 and 17 (73)

## Key

- Exp    explode subject heading (searches additional, closely related but more specific terms)  
/       Medical Subject Heading (MeSH)  
\*       truncated term identifying terms with the same stem  
" ... "    searches for the two or more words within the quotation marks as a phrase  
.ti      searches title  
Or/1–3 set combination 1 or 2 or 3

**Fig. 2.3** Example search for incidence and prevalence of hearing loss in the UK

required for the model. However, where there are few or perhaps no RCTs, a broader search for other study types is essential. Also, data from RCTs may not be appropriate to use if the population or condition differs to that in your model. Other sources of evidence of effectiveness such as observation studies (e.g. cohort studies) are important in providing the model with evidence from the ‘real world’ or more directly relevant information. Initial searches for systematic reviews and RCTs of clinical effectiveness can be supplemented by further targeted searches for observational studies.

Systematic reviews, trials and other evaluation studies of the effectiveness and safety of an intervention can be found in journal papers, conference abstracts, theses and other unpublished reports. Searching MEDLINE alone is not sufficient for identifying trials for a systematic review (Royle et al. 2005; Lefebvre et al. 2011) and should not be relied upon to provide an unbiased, generalisable set of effectiveness studies. MEDLINE, Embase and the Cochrane Library are recommended as a minimum set of databases (Lefebvre et al. 2011) by the Cochrane Collaboration. Similarly the Centre for Reviews and Dissemination recommends searching MEDLINE, Embase and further relevant databases to the questions (such as PsycINFO for evidence on mental health interventions) (Centre for Reviews and

**Table 2.5** Selected clinical effectiveness resources

Resource	Study type	URL (accessed 29-9-14)
Cochrane Database of Systematic Reviews	Systematic reviews	<a href="http://www.thecochranelibrary.com/">www.thecochranelibrary.com/</a>
Database of Abstracts of Reviews of Effectiveness	Systematic reviews	<a href="http://www.thecochranelibrary.com/">www.thecochranelibrary.com/</a>
Health Technology Assessments	Reviews and syntheses	<a href="http://www.thecochranelibrary.com/">www.thecochranelibrary.com/</a>
NHSEED	Economic evaluations	<a href="http://www.thecochranelibrary.com/">www.thecochranelibrary.com/</a>
Cochrane Central Register of Controlled Trials	Controlled trials	<a href="http://www.thecochranelibrary.com/">www.thecochranelibrary.com/</a>
MEDLINE/PubMed	Trials, reviews and other evaluative studies	<a href="http://www.ncbi.nlm.nih.gov/pubmed">www.ncbi.nlm.nih.gov/pubmed</a>
Embase	Trials, reviews and other evaluative studies	<a href="http://www.elsevier.com/online-tools/Embase">www.elsevier.com/online-tools/Embase</a>
NHS Evidence	Trials, reviews and other evaluative studies	<a href="http://www.evidence.nhs.uk">www.evidence.nhs.uk</a>
McMaster Health Systems Evidence	Reviews and syntheses	<a href="http://www.mcmasterhealthforum.org/hse/">www.mcmasterhealthforum.org/hse/</a>
Trip Database	Trials, reviews and other evaluative studies	<a href="http://www.tripdatabase.com/">http://www.tripdatabase.com/</a>

Dissemination 2011). A combination of searching MEDLINE, Embase and NHSEED found 87.3 % of the clinical effectiveness studies in technology appraisal reports (TARs) (Royle and Waugh 2003). For quicker retrieval of evidence, Trip and NHS Evidence are user-friendly web portals of selected evidence-based health (EBH) studies and reports. Many studies are also found in MEDLINE, Embase and the Cochrane Library, but for the purpose of populating a model, they offer rapid access to key reports which may be sufficient. Selected key health databases for identifying RCTs, systematic reviews and other evaluative studies are listed in Table 2.5. Reports of trials have been found in many national and international databases of both published and unpublished literature (AUHE 2015).

The search strategy to identify evidence of effectiveness is likely to include search concepts for the ‘Patient/Population’ and ‘Intervention’ under consideration. A further search concept can be added to limit the results to effectiveness studies, e.g. RCTs, systematic reviews and observational studies. This third concept can be added to the search either by using a search filter or a database limit or developing a search strategy for the concept. The database limits are very quick to apply and should retrieve a small number of highly relevant results, but they will miss relevant studies that are found when using the search filter.

Where the size of available evidence is small, then a search of just ‘population’ and ‘intervention’ without a limit to study type will suffice. The analyst can view a

**Table 2.6** Search concepts and terms for effectiveness studies of Herceptin use in breast cancer

Search concept	(A) Breast cancer		(B) Herceptin	(C) Effectiveness studies
Search terms	Breast	Cancer*	Trastuzumab	Randomised controlled trial*
	Mammary	Neoplasm*	Herceptin	Placebo
		Tumor*	Herclon	Double blind*
		Carcinoma*		Review*
		Adenocarcinoma*		Meta-analysis*
		Malignant*		Observational
		Sarcoma*		Cohort*

wider set of evidence, of varying quality, and select the best quality and most appropriate data for their model. For example, the question ‘How effective is Herceptin in breast cancer?’ contains the search concepts (A) ‘breast cancer’, (B) ‘Herceptin’ and (C) ‘effectiveness studies’. If relevant trials were found, then the appropriate search combination would be A AND B AND C to retrieve records mentioning a breast cancer term, a Herceptin term and a trials term. If there were no trials, then a simpler search A AND B would identify all studies with the terms breast cancer and Herceptin, from which the analyst chooses the most appropriate to inform the model. Table 2.6 illustrates possible search concepts and suggested terms. Please note that there are clinical effectiveness search filters available for some databases available to use as an alternative to the information specialist/analyst creating their own clinical effectiveness search.

A search can be limited to only retrieve particular study types, for example, RCTs, by using a tested search filter. This replaces the need to identify and test your own collection of terms for an RCT search. The Cochrane Handbook searching for studies chapter includes validated RCT search filters designed for use in Ovid MEDLINE and PubMed that can be copied and pasted (Lefebvre et al. 2011). The Cochrane RCT filter (precision maximising) has been used to identify clinical effectiveness studies (RCTs) of Herceptin in breast cancer in Fig. 2.4 lines 9–18. Use of a search filter within your search should be acknowledged and documented in the search methods section of the final report. A collection of ‘study type’ filters is available from the InterTASC ISSG Search Filters Resource webpage (InterTASC ISSG 2015a). The filters have been published on websites or in journal papers. Some have been validated, and some have been peer reviewed, but not all. The filters for RCTs, systematic reviews, observational studies and therapy studies may be helpful for identifying effectiveness studies. Before using a filter it is important to note which database it has been designed for. A search strategy designed for MEDLINE will have less effective retrieval in Embase than in MEDLINE due to the differences in indexing terms used (MeSH in MEDLINE, EMTREE in Embase). Filters should be used in the appropriate database or adapted to take account of different indexing terms.

Ovid MEDLINE(R) <1946 to September Week 1 2014>Search Strategy:

- 
- 1 exp breast neoplasms/ (222893)
  - 2 (breast adj5 (neoplasm\* or cancer\* or tumour\* or tumor\* or carcinoma\* or adenocarcinoma\* or malignant\* or sarcoma\*)).ti,ab.(214346)
  - 3 (mammary adj5 (neoplasm\* or cancer\* or tumour\* or tumor\* or carcinoma\* or adenocarcinoma\* or malignant\* or sarcoma\*)).ti,ab.(26991)
  - 4 or/1–3 [Breast Cancer] (281799)
  - 5 trastuzumab.ti,ab.(4569)
  - 6 herceptin.ti,ab.(1377)
  - 7 herclon.ti,ab.(0)
  - 8 or/5–7 [Trastuzumab] (5209)
  - 9 randomized controlled trial.pt. (387973)
  - 10 controlled clinical trial.pt. (89778)
  - 11 randomized.ab. (283989)
  - 12 placebo.ab. (150625)
  - 13 clinical trials as topic/ (173007)
  - 14 randomly.ab. (200735)
  - 15 trial.ti. (123439)
  - 16 9 or 10 or 11 or 12 or 13 or 14 or 15 (885427)
  - 17 exp animals/ not humans/. (4009223)
  - 18 16 not 17 [Cochrane RCT filter -precision maximising] (813525)
  - 19 4 and 8 and 18 (835)

#### Key

- Exp    explode subject heading (searches additional, closely related but more specific terms)
- /       Medical Subject Heading (MeSH)
- Adj 5   terms must be within 5 words of each other
- \*       truncated term identifying terms with the same stem
- “ ... ”   searches for the two or more words within the quotation marks as a phrase
- .ti, ab   searches title and abstract
- .ti      searches title
- .ab     searches abstract
- .pt    searches publication type
- Or/1–3 set combination 1 or 2 or 3

**Fig. 2.4** Clinical effectiveness of Herceptin for breast cancer search using Cochrane RCT filter (precision maximising)

### 2.5.3 Database Limits and Clinical Queries

Ovid databases have an ‘additional limits’ feature containing the publication-type limit ‘randomised controlled trial’ and the clinical query limits for ‘reviews’ and ‘therapy’ studies. These can be used to identify effectiveness studies in Ovid MEDLINE. Lines 10 and 11 are clinical queries, whereas line 12 is a publication type limit in Fig. 2.5. Similar queries are available in PubMed. EBSCOhost databases (e.g. CINAHL) also have database limit options. NHS Evidence has filters for ‘types of information’ that can be used to limit a search to effectiveness studies. The ‘primary research’ and ‘health technology assessments’ can identify effectiveness studies.

Ovid MEDLINE(R) <1946 to September Week 1 2014>

Search Strategy:

- 
- 1 exp breast neoplasms/ (222893)
  - 2 (breast adj5 (neoplasm\* or cancer\* or tumour\* or tumor\* or carcinoma\* or adenocarcinoma\* or malignan\* or sarcoma\*)).ti,ab.(214346)
  - 3 (mammary adj5 (neoplasm\* or cancer\* or tumour\* or tumor\* or carcinoma\* or adenocarcinoma\* or malignan\* or sarcoma\*)).ti,ab.(26991)
  - 4 or/1–3 [Breast Cancer] (281799)
  - 5 trastuzumab.ti,ab.(4569)
  - 6 herceptin.ti,ab.(1377)
  - 7 herclon.ti,ab.(0)
  - 8 or/5–7 [Trastuzumab] (5209)
  - 9 4 and 8 (4162)
  - 10 limit 9 to "reviews (maximizes specificity)" (90)
  - 11 limit 9 to "therapy (maximizes specificity)" (184)
  - 12 limit 9 to randomized controlled trial (182)

Key

Exp explode subject heading (searches additional, closely related but more specific terms)

/ Medical Subject Heading (MeSH)

Adj 5 terms must bewithin 5 words of each other

\* truncated term identifying terms with the same stem

.ti, absearches title and abstract

Or/1–3 set combination 1 or 2 or 3

**Fig. 2.5** Clinical effectiveness of Herceptin for breast cancer search using Ovid ‘Additional Limits – Clinical Queries’

Adverse events data for a cost effectiveness model could include side effects of the health intervention, complications, treatment failure and safety issues. Estimates of cost effectiveness could be biased if data for costs, incidence and consequences of adverse events are absent (Heather et al. 2014). Adverse events may be reported in the trial and systematic reviews that contribute clinical outcomes. However, follow-up studies are more likely to have had a suitable length of time to collect adverse events data. The clinical effectiveness resources listed in Table 2.5 and outlined in Sect. 2.5.2 are applicable to adverse events data. Recent evaluations of literature search methods for adverse effects reviews indicate that searching MEDLINE and Embase with adverse effects search filters and terms can retrieve the majority of studies (Golder et al. 2014). Other important sources include industry submissions, reference lists and (for drug safety) databases such as TOXLINE and Derwent Drug Index (Golder and Loke 2012b). Subject-specific databases could also prove useful.

Developing an effective search strategy for identifying adverse events can be problematic due to the variety of terms and synonyms that indicate adverse events. These can be generic, for example, adverse effects, complications, harms, risks, safety, side effects and toxicity, or can be specific, for example, rash, pain and nausea. It is advisable to seek advice from a librarian or information specialist when developing an adverse event search since this is an active research area.

An adverse events search should include the search concept for the intervention under consideration and may include a Patient/Population search concept (if this is appropriate to the model). A further search concept would be added to limit the search to studies mentioning adverse events. If using Medical Subject Headings (MeSH) in MEDLINE, the adverse events subheading can be used to limit the search to studies about adverse events associated with the intervention (Golder and Loke 2012a). For example, using the MeSH ‘Hearing Aids/ae [Adverse Effects]’ will retrieve studies of hearing aid complications and adverse events. Other options include using an adverse effects search filter (several are available via the ISSG Search Filter Resource (InterTASC ISSG 2015b), using the adverse effects floating subheading or developing a search strategy using adverse event terms appropriate to your research question. NHS Evidence has a filter for limiting results to ‘drug prescribing and safety’ studies.

## 2.6 Finding the Evidence on Health-Related Quality of Life and Health State Preferences

In Chap. 1 (Section 1.3.3) we described how within CUA the effect is measured in terms of ‘healthy years’ which are represented by a multidimensional utility-based measure (where a utility is a measure of preference) which combines life years gained with some judgement on the quality of those life years. When searching for evidence on health-related quality of life (HRQoL), there are generic preference-based measures such as the EQ-5D (EuroQol 1990), SF-6D (Brazier et al. 2002), HUI2 and HUI3 (Feeny et al. 2002) which can be used across different conditions and disease areas. In addition condition-specific utility measures have been developed with the aim of providing a more accurate assessment of the impact of conditions and posited to provide a more sensitive measure of the benefit of interventions; for example, measures have been developed for pressure ulcers (Czoski-Murray et al. 2013) and incontinence (Brazier et al. 2008). When building a cost effectiveness model, the analyst will invariably need to find evidence on HRQoL. However, in some cases evidence from preference-based measures may not be available, or those that are available may be inappropriate. In these cases the analyst may look more widely at non-preference-based measures. For example, Fairburn et al., in their model of the cost effectiveness of transcatheter aortic valve implantation versus surgical aortic valve replacement in patients with severe aortic stenosis at high operative risk, use the New York Heart Association (NYHA) class transitions with each class ascribed a mean EQ-5D utility value (Fairbairn et al. 2013).

### 2.6.1 Where to Look

Health utilities, health state preferences, values and HRQoL data are available in journal articles and via databases or registries of instruments and preferences. Sources include general health databases, speciality databases and instrument-specific

**Table 2.7** Selected sources of reports of health utilities, HRQoL and health state preferences

Resource	Type	URL (access 2-10-14)
ScHARRHUD	Studies of health state utility values	<a href="http://www.scharrhud.org">www.scharrhud.org</a>
PROQOLID	Patient-reported outcome and quality of life instruments	<a href="http://www.proqolid.org/">http://www.proqolid.org/</a>
MEDLINE	General health database with greater N American journal coverage	<a href="http://www.ncbi.nlm.nih.gov/pubmed">www.ncbi.nlm.nih.gov/pubmed</a>
Embase	General health database with greater European journal coverage	<a href="http://www.elsevier.com/online-tools/Embase">www.elsevier.com/online-tools/Embase</a>
NHSEED	Economic evaluations, UK focussed	<a href="http://www.crd.york.ac.uk/CRDWeb/">www.crd.york.ac.uk/CRDWeb/</a>
RePEC	Published and working papers in economics	<a href="http://repec.org/">http://repec.org/</a>
CEA Registry	CEA studies	<a href="https://research.tufts-nemc.org/cear4/">https://research.tufts-nemc.org/cear4/</a>
EQ-5D	Individual instrument website	<a href="http://www.euroqol.org/">http://www.euroqol.org/</a>
SF-36	Individual instrument website	<a href="http://www.sf-36.org/">http://www.sf-36.org/</a>

websites. Table 2.7 offers a brief list of selected sources, but more a comprehensive list is available from the Etext on Health Technology Assessment Information Resources (Paisley et al. 2005). Subject-specific databases may also be worth searching for appropriate health utility data.

### 2.6.2 *Search Strategy, Concepts, Terms and Combinations*

The Intervention and Patient/Population are likely to form two search concepts (similar to the clinical effectiveness, epidemiology and the existing model search). A further concept is required to limit the search to studies that mention HRQoL or health state preference terms. Generic and/or specific terms for HRQoL measures and health state preferences can be used. The choice of terms can affect the numbers retrieved and the specificity of the search greatly. Table 2.8 lists example terms under their search concept. These terms have been suggested by health economists, identified through textual analysis with some selections from the Etext on Health Technology Assessment Information Resources (Paisley et al. 2005). It is by no means a complete listing, and there are many more specific and generic instrument terms that could be included in a search as appropriate.

For example, a search for studies mentioning HRQoL or health state preferences in breast cancer patients using Herceptin would be executed by combining the search concepts using ‘AND’, i.e. A ‘breast cancer’ AND B herceptin AND C

**Table 2.8** Search concepts and terms for HRQoL and health state preferences in Herceptin use in breast cancer

Search concept	(A) Breast cancer		(B) Herceptin	(C) HRQoL/health state preferences
Search terms	Breast	Cancer*	Trastuzumab	Health utilities
	Mammary	Neoplasm*	Herceptin	Health-related quality of life
		Tumor*	Herclon	Preference weights
		Carcinoma*		Preference scores
		Adenocarcinoma*		Health status
		Malignan*		Instrument scores
		Sarcoma*		Questionnaire
				QALY
				HUI
				EQ-5D
				QWB
				SF-36

HRQoL or health state preferences). This combination is illustrated in Fig. 2.6, search line 29. The following search line (30) indicates how limiting a search to specific measures reduced the number of records retrieved. The search line (29) which includes the generic terms identifies far more records (142) when compared with two records found in the specific-measures-only search.

### 2.6.3 *Search Filters, Database Limits and Clinical Queries*

There are currently no validated search filters for health utilities; however, the quality of life studies section of the ISSG Search Filters Resource offers guidance on searching HRQoL studies (Paisley et al. 2005), outcome studies (Brettle et al. 1998) and properties of measurement instruments (Terwee et al. 2009). Databases do not contain ready made ‘clinical queries’ or other options for quickly limiting a search to HRQoL or health state preference studies.

## 2.7 Finding Evidence on Resource Use and Costs

For any economic evaluation it is important to determine resource use and then assign a unit cost to it. We talk more about costs in Chaps. 3 and 7, but in essence the resources included will depend on the perspective the evaluation is taking; this in turn will be informed by the question being asked and the purpose of the analysis. The costs might include costs to the health care system, costs to

Ovid MEDLINE(R) <1946 to September Week 4 2014>

Search Strategy:

- 
- 1    exp breast neoplasms/ (223494)
  - 2    (breast adj5 (neoplasm\* or cancer\* or tumour\* or tumor\* or carcinoma\* or adenocarcinoma\* or malignant\* or sarcoma\*)).ti,ab.(215009)
  - 3    (mammary adj5 (neoplasm\* or cancer\* or tumour\* or tumor\* or carcinoma\* or adenocarcinoma\* or malignant\* or sarcoma\*)).ti,ab.(27024)
  - 4    or/1-3 [Breast Cancer] (282567)
  - 5    trastuzumab.ti,ab.(4601)
  - 6    herceptin.ti,ab.(1384)
  - 7    herclon.ti,ab.(0)
  - 8    or/5-7 [Herceptin] (5246)
  - 9    Cost-Benefit Analysis/ (61689)
  - 10    health status indicators/ (20462)
  - 11    "Quality of Life"/ (122361)
  - 12    Health status/ (61194)
  - 13    questionnaire/ (312051)
  - 14    cross-sectional studies/ (185868)
  - 15    ("Health related quality of life" or hqol or h qol or hrqol or hr qol or pqol or qls).ti,ab.(21759)
  - 16    quality adjusted life years/ (7352)
  - 17    ("quality adjusted life" or qaly or qalys or qald or qale or qtime).ti,ab.(7100)
  - 18    (utility\* adj5 (health\* or score\* or weight\*)).ti,ab.(3252)
  - 19    (preference\* adj5 (health\* or score\* or weight\*)).ti,ab.(3490)
  - 20    (instrument\* adj5 (health\* or score\* or weight\*)).ti,ab.(6516)
  - 21    (HYE or HYES or "health\* year\* equivalent").ti,ab.(63)
  - 22    health state.ti,ab.(2488)
  - 23    or/9-22 [Generic Terms] (660974)
  - 24    (hui1 or hui2 or hui3).ti,ab.(277)
  - 25    ("quality of wellbeing" or "quality of well being" or qwb).ti,ab.(364)
  - 26    (eq-5d or eq5d or euroqol\*).ti,ab.(2959)
  - 27    or/24-26 [Specific Measures] (3480)
  - 28    23 or 27 (661184)
  - 29    4 and 8 and 28 [Generic or Specific Health Utilities in Breast Cancer Herceptin Studies] (142)
  - 30    4 and 8 and 27 [Specific-only Health Utilities in Breast Cancer Herceptin Studies] (2)

**Key**

Exp    explode subject heading (searches additional, closely related but more specific terms)

/       Medical Subject Heading (MeSH)

Adj 5    terms must be within 5 words of each other

\*       truncated term identifying terms with the same stem

.ti, ab    searches title and abstract

Or/1-3    set combination 1 or 2 or 3

**Fig. 2.6** HRQoL and health state preferences in studies of Herceptin and breast cancer

the economy in general or even costs tailored to the specifications of a particular decision maker. For the analyst building a cost effectiveness model, evidence relating to resource use and costs should be scrutinised closely. The analyst should consider the generalisability of resource use and costs between settings and inadequacy of exchange rate translation as a means for estimating national costs based upon data from other settings. Clinical guidelines from health systems and professional bodies are a source of resource use recommendations,

although often they do not have cost data associated with the resource use. As a rule of thumb, there is a general health warning in using cost data from previously published literature – although if that is all we have, then it is better than nothing.

As costs are unavoidably provider specific, evaluations for single payer health systems will often be easier than for more distributed health systems. Hence evaluations in the UK can refer to the NHS reference costs (UK Department of Health 2013) for hospital-based costs and to the PSSRU unit costs for health and social care (PSSRU 2014). By contrast, economic evaluations undertaken in North America will need to draw upon data from the public, private and not-for-profit sectors. For example, in the USA the physician fee component of cost can be obtained from Medicare (CMS 2014); however, hospital, test and pharmaceutical costs will often need to be obtained separately and frequently vary according to provider and indeed the payer (Robinson 2011). In the USA, Medicare provides increasingly detailed information on costs they incur for a wide range of services (CMS 2012). In Canada the Canadian Institute for Health Information (CIHI) has produced a costing tool, called the Patient Cost Estimator that provides province-specific costs for hospital procedures based upon an age group-specific case mix group system (CIHI 2014).

### ***2.7.1 Where to Look***

Resource use and cost data are available from databases of national reference costs and health care use. Literature searches can also identify economic evaluations and cost studies that may contain relevant cost data. A selected list of data sources and research databases is offered in Table 2.9.

### ***2.7.2 Search Strategy, Concepts, Terms and Combinations***

The first step is to identify costs from reference sources, for example, NHS reference costs. If this fails to identify sufficient costs data, then a search for economic evaluations that may contain relevant cost data can be undertaken. The search strategy usually comprises three or four concepts: firstly the intervention you require costs for and secondly the Patient/Population who would use the resource. A further search concept for economic evaluations or cost studies is added if you are searching a general database. Figure 2.1 is an example of a targeted search for economic evaluations in MEDLINE using a limited number of search terms to identify economic evaluation terms. Other search filters and clinical queries are available and can be used instead of the terms suggested in Figure 2.1 for more sensitive searches.

**Table 2.9** Selected sources of reports on cost, utilities, effectiveness and safety

Resource	Type	URL (access 2-10-14)
BNF Drug Prices	British National Formulary provided dosage, cost and safety information for drugs licensed for use in the UK NHS	<a href="https://www.medicinescomplete.com/mc/bnf/current/">https://www.medicinescomplete.com/mc/bnf/current/</a>
Canadian Institute for Health Information	Cost data for hospital procedures in Canada	<a href="http://www.cihi.ca/CIHI-ext-portal/internet/en/documentfull/spending+and+health+workforce/spending/pce_application">http://www.cihi.ca/CIHI-ext-portal/internet/en/documentfull/spending+and+health+workforce/spending/pce_application</a>
Centre for Medicare and Medicaid	Cost data for health care providers in the USA	<a href="http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Medicare-Provider-Charge-Data/index.html">http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Medicare-Provider-Charge-Data/index.html</a>
ClinicalTrials.gov	Registry and results database of clinical studies	<a href="https://clinicaltrials.gov">https://clinicaltrials.gov</a>
Cost effectiveness Registry	Annotated database of published cost effectiveness analyses	<a href="https://research.tufts-nemc.org/cear4/SearchtheCEARegistry.aspx">https://research.tufts-nemc.org/cear4/SearchtheCEARegistry.aspx</a>
Embase		<a href="http://www.elsevier.com/online-tools/Embase">www.elsevier.com/online-tools/Embase</a>
Expert opinion		<a href="http://www.crd.york.ac.uk/CRDWeb/">www.crd.york.ac.uk/CRDWeb/</a>
Health Technology Assessments		<a href="http://www.ncbi.nlm.nih.gov/pubmed">www.ncbi.nlm.nih.gov/pubmed</a>
MEDLINE	Summary of the UK NHS foundation trust DRG costs	<a href="https://www.gov.uk/government/collections/nhs-reference-costs">https://www.gov.uk/government/collections/nhs-reference-costs</a>
NHS reference costs	Fees paid to NHS foundation trusts by NHS England	<a href="https://www.gov.uk/government/publications/national-tariff-payment-system-2014-to-2015">https://www.gov.uk/government/publications/national-tariff-payment-system-2014-to-2015</a>
NHSEED		<a href="http://www.crd.york.ac.uk/CRDWeb/">www.crd.york.ac.uk/CRDWeb/</a>
PSRRU unit costs of health and social care	Bottom-up unit cost estimates for the UK social and NHS community care services	<a href="http://www.psrn.ac.uk/project-pages/unit-costs/">http://www.psrn.ac.uk/project-pages/unit-costs/</a>
Registry data		<a href="https://patientregistry.ahrq.gov">https://patientregistry.ahrq.gov</a>
Technology manufacturers		

### 2.7.3 *Search Filters, Database Limits and Clinical Queries*

Several search filters have been developed to identify economic evaluations. There is no single recommended filter; however, the NHS CRD NHSEED search strategy (Centre for Reviews and Dissemination 2014) is often used as a sensitive search for economic evaluations. Three cost and three economics search filters developed by Wilczynski et al. (2004) are available as ‘clinical queries’ within MEDLINE, PubMed and Embase. These can be added to an existing search strategy to limit the results to costs or economic studies.

## 2.8 Tracking and Reporting Search Activities

The methods used to identify data for a model plus the justification for selecting particular data should be documented in reports of cost effectiveness models. A transparent account of sources used, search strategies and other approaches used to identify data enables readers to judge the currency and relevancy of the information. Poor reporting, where there is unclear, incomplete or misleading information in the study, can lead to poor and costly decision-making (Husereau et al. 2013). The Consolidated Health Economic Evaluation Reporting Standard (CHEERS) statement provides a checklist to guide reporting economic evaluations. Item 13b suggests that model-based economic evaluations ‘describe approaches and data sources used to estimate resource use associated with model health states. Describe primary or secondary research methods for valuing each resource item in terms of its unit cost. Describe any adjustments made to approximate opportunity costs’ as a bare minimum (Husereau et al. 2013).

Tracking search activities throughout the model development will help collate the information needed to accurately describe the search approaches and data sources used in the model. This can be a considerable task since there are likely to be multiple searches and many different resources consulted. Maintaining a joint search activity log for both information specialists and analysts helps track all search activities during the model’s development. Table 2.10 offers an example of a search activity log capturing full details of the resources, dates of searches (or requests), purpose of search and use in model, results of the search and notes to capture important issues for the search activity. It is good practice to save search strategies that have been (or are being) developed in a document and in a database account. This allows searches to be rerun or edited at a later date and provides a record of exactly which terms have been used. Some reports (e.g. health technology assessments) require copies of search strategies used. The combination of an up-to-date search activity log, copies of all search strategies used and any documentation for obtaining data directly from sources other than research databases will allow a full and transparent account of the search approach and data sources.

**Table 2.10** Search activity log

Resource	Date of search/contact	Purpose/use in model	Search results	Notes
Ovid MEDLINE(R) <1946 to September Week 1 2014>	10-09-14	RCT search for clinical effectiveness data	147 hits	No date or language limits applied
Ovid MEDLINE(R) <1946 to September Week 1 2014>	13-09-14	Health utilities	56 hits	
NHSEED (Wiley)	05-08-14	Identify existing models to aid model design	35 hits	
OECD Health Statistics <a href="http://stats.oecd.org/">http://stats.oecd.org/</a>	20-10-14	Disease incidence and prevalence figures	1 report	This report was selected as it contained more current data than others
PSSRU unit costs of health and social care <a href="http://www.pssru.ac.uk/project-pages/unit-costs/">http://www.pssru.ac.uk/project-pages/unit-costs/</a>	27-10-14	Resource costs		
Industry contact	27-10-14	Resource use	Awaiting response	Emailed expert for further data

## 2.9 Quality Assessment Tools

Standard hierarchies of evidence tend not to be directly applicable to the assessment of the quality of evidence used in cost effectiveness models. The overarching principle is that the best available evidence should be used; however, assessing what is the best evidence to inform a decision involves the balancing of the quality of the evidence with its relevance to the decision problem. Is it better to have perfect evidence of the wrong parameter or imperfect evidence on the right parameter? Ideally of course, we want perfect evidence on the correct parameter, but this is rarely available. Quality assessment for the evidence used in cost effectiveness models often combines standardised tools such as GRADE (GRADE Working Group 2014), QUADAS-2 (Whiting et al. 2011) and CHARMS (Moons et al. 2014) to support judgements about whether the highest quality relevant evidence was used (Caro et al. 2012). Hence much of the focus to promote quality improvement in cost effectiveness modelling has focussed on transparent reporting of the processes used to identify the evidence used in models and the rationale for choosing specific evidence when alternatives were available (Husereau et al. 2013).

## 2.10 Summary

- Evidence used in models of cost effectiveness comes from many sources including evidence syntheses, expert judgement, observational studies, RCTs, other clinical studies, reference sources and routine data sources.
- The two key databases to search are MEDLINE and Embase.
- There are standard approaches to developing a literature search that requires identifying words, phrases and index terms (database subject headings) which are present in studies relevant to the research question. The results of the searches are then combined using ‘Boolean’ logic to produce a final set of records of references that are relevant to the question.
- A targeted literature search to identify existing cost effectiveness models at the start of the modelling process can help avoid duplicating work previously undertaken.
- Clinical evidence or information is required to accurately predict the incidence and describe the prevalence of a disease in a given population within the model. Evidence on natural history of a disease provides a view over time of what resources may be called upon and can indicate alternative care pathways.
- Evidence on health-related quality of life may take the form of generic preference-based measures such as the EQ-5D, SF-6D, HUI2 and HUI3 or condition-specific utility measures.
- Evidence relating to resource use and costs should be scrutinised closely. The analyst should consider the generalisability of resource use and costs between settings and inadequacy of exchange rate translation as a means for estimating national costs based upon data from other settings.
- The methods used to identify data plus the justification for selecting particular data should be documented. A transparent account of sources used, search strategies and other approaches used to identify data enables readers to judge the currency and relevancy of the information.

## References

- Academic Unit of Health Economics (2015) Finding randomised controlled trials. [http://med-health.leeds.ac.uk/info/639/information\\_specialists/1790/finding\\_randomised\\_controlled\\_trials](http://med-health.leeds.ac.uk/info/639/information_specialists/1790/finding_randomised_controlled_trials) (accessed 26-06-15)
- Academy Health (2011) Health Economics Core Library Recommendations. <http://www.nlm.nih.gov/nichsr/corelib/hecon-2011.html> (accessed 07-04-2015)
- Aveyard H (2010) Doing a literature review in health and social care: a practical guide: a practical guide. McGraw-Hill/Open University Press, Maidenhead
- BIREME (2015) LILACs Latin American and Caribbean Center on Health Sciences Information <http://lilacs.bvsalud.org/en/> (accessed 26-06-15)
- Booth A, Diana P, Anthea S (2011) Systematic approaches to a successful literature review. Sage, Thousand Oaks

- Brazier J, Roberts J, Deverill M (2002) The estimation a preference-based single index measure for health from the SF-36. *J Health Econ* 21(2):271–292
- Brazier JE, Czoski-Murray C, Roberts J, Brown M, Symonds T, Kelleher C (2008) Estimation of a preference-based index from a condition-specific measure: the King's health questionnaire. *Med Decis Making* 28(1):113–126
- Brettle AJ et al (1998) "Searching for information on outcomes: do you need to be comprehensive?" *Quality in Health Care* 7.3:163–167
- Caro J, Briggs AH, Siebert U, Kuntz KM (2012) Modelling research good practices – overview: a report of the ISPOR-SMDM modelling good research practices task force – 1. *Med Decis Making* 32(5):667–677
- Centre for Reviews and Dissemination (2011) Systematic Reviews: CRD's Guidance for Undertaking Reviews in Healthcare [Internet]. Centre for Reviews & Dissemination, York. [cited 2012 Dec 7]. p 281. Available from: [http://www.york.ac.uk/inst/crd/pdf/Systematic\\_Reviews.pdf](http://www.york.ac.uk/inst/crd/pdf/Systematic_Reviews.pdf)
- Center for Reviews and Dissemination (2014) Search Strategies. <http://www.crd.york.ac.uk/CRDWeb/searchstrategies.asp> (accessed 07-4-15)
- CIHI (2014) Canadian Institute for Health Information. [http://www.cihi.ca/CIHI-ext-portal/internet/en/documentfull/spending+and+health+workforce/spending/pce\\_application](http://www.cihi.ca/CIHI-ext-portal/internet/en/documentfull/spending+and+health+workforce/spending/pce_application) (accessed 13-4-15)
- CMS (2012) Medicare Provider Utilization and Payment Data. <https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Medicare-Provider-Charge-Data/index.html> (accessed 07-4-15)
- CMS (2014) Physician fee schedule overview. <https://www.cms.gov/apps/physician-fee-schedule/overview.aspx>
- Czoski-Murray C, Meads D, Edlin R, Hulme CT, Gorecki C, Nixon J, McCabe C Constructing a utility algorithm for the pressure ulcer quality of life-utility instrument (PUQOL-UI) working paper WP14-04, Academic Unit of Health Economics, Leeds Institute of Health Sciences, University of Leeds
- EuroQol Group (1990) EuroQol-a new facility for the measurement of health-related quality of life. *Health Policy* 16(3):199–208
- Fairbairn T, Meads D, Hulme C, Mather A, Blackman D, Plein S, Greenwood J (2013) The cost-effectiveness of transcatheter aortic valve implantation versus surgical aortic valve replacement in patients with severe aortic stenosis at high operative risk. *Heart* 99:914–920. doi:[10.1136/heartjnl-2013-303722](https://doi.org/10.1136/heartjnl-2013-303722)
- Feeny D, Furlong W, Torrance GW, Zenglong Z, DePauw S et al (2002) *Med Care* 40(2):113–128
- Glanville J, Paisley S (2010) Searching for cost-effectiveness decisions. In: Shemilt I, Mugford M, Vale L, Marsh K, Donaldson C (eds) Evidence-based decisions and economics. Wiley-Blackwell, Oxford, pp 79–92
- Golder S, Loke YK (2012a) Failure or success of electronic search strategies to identify adverse effects data. *J Med Libr Assoc* 100(2):130–134
- Golder S, Loke YK (2012b) The contribution of different information sources for adverse effects data. *Int J Technol Assess Health Care* 28(2):133–137
- Golder S, Loke YK, Zorzela L (2014) Comparison of search strategies in systematic reviews of adverse effects to other systematic reviews. *Health Inf Libr J* 31(2):92–105
- GRADE Working Group (2014) <http://www.gradeworkinggroup.org/index.htm>
- Heather EM, Payne K, Harrison M, Symmons DP (2014) Including adverse drug events in economic evaluations of anti-tumour necrosis factor- $\alpha$  drugs for adult rheumatoid arthritis: a systematic review of economic decision analytic models. *Pharmacoeconomics* 32(2):109–134
- Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, Augustovski F, Briggs AH, Mauskopf J, Loder E (2013) Consolidated health economic evaluation reporting standards (CHEERS) statement. *BMC Med* 11(1):80
- InterTASC ISSG (2015a) The InterTASC Information Specialists' Sub-Group Search Filter Resource. <https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home>. (accessed 07-04-15)

- InterTASC ISSG (2015b) ISSG Search Filter Resource; Adverse Events Filters. <https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/adverse-events-filters> (accessed 07-04-15)
- Johnston K, Buxton MJ, Jones DR, Fitzpatrick R (1999) Assessing the costs of healthcare technologies in clinical trials. *Health Technol Assess* 3(6):1–76
- Kalenthaler E, Essat M, Tappenden P, Paisley S (2014) The identification and review of cost effectiveness model parameters: a qualitative study. *Int J Technol Assess Health Care* 30(3):333–40
- Lefebvre C, Manheimer E, Glanville J (2011) Chapter 6: Searching for studies. In: Higgins JPT, Green S (eds) *Cochrane Handbook for Systematic Reviews of Interventions* version 5.1.0 (updated March 2011). The Cochrane Collaboration
- McMaster University (2015) Health Systems Evidence. <http://www.mcmasterhealthforum.org/hse> (accessed 07-04-15)
- Moons KGM, de Groot JHA, Bouwmeester W, Vergouwe Y, Mallett S et al (2014) Critical appraisal and data extraction for the systematic reviews of prediction modelling studies: the CHARMS Checklist. *PLoS Med* 11(10): <http://www.plosmedicine.org/article/fetchObject.action?uri=info:doi:10.1371/journal.pmed.1001744&representation=PDF>
- National Institute for Health and Care Excellence (2013) Guide to the methods of technology appraisal. London
- National Library of Medicine (2014) PubMed. [http://www.nlm.nih.gov/pubs/factsheets/dif\\_med\\_pub.html](http://www.nlm.nih.gov/pubs/factsheets/dif_med_pub.html)
- Paisley S (2010) The Classification of evidence in decision-analytic models of cost-effectiveness: a content analysis of published report. *Int J Technol Assess Health Care* 26(4):458–462
- Paisley S, Booth A, Mensinkai S (2005) Chapter 12: Health-related quality of life studies. Etext on Health Technology Assessment (HTA) Information Resources. <http://www.nlm.nih.gov/archive/20060905/nichsr/ehta/chapter12.html>
- PSSRU (2014) Unit costs of health and social care. University of Kent, Canterbury. <http://www.pssru.ac.uk/project-pages/unit-costs/>
- QUADAS (2014) A quality assessment tool for diagnostic accuracy studies. <http://www.bris.ac.uk/quadas/>
- Robinson JC (2011) Variation in hospital costs, payments and profitability for Cardiac Valve Replacement Surgery. *Health Serv Res* 46(6:1):1928–1945
- Royle P, Waugh N (2003) Literature searching for clinical and cost-effectiveness studies used in health technology assessment reports carried out for the National Institute for Clinical Excellence appraisal system. *Health Technol Assess* 7(34)
- Royle P, Bain L, Waugh N (2005) Systematic reviews of epidemiology in diabetes: finding the evidence. *BMC Med Res Methodol* 5(1):2
- Sackett DL, Richardson WS, Rosenberg W, Haynes RB (1997) Evidence-based medicine: how to practice and teach EBM. Churchill Livingston, New York
- Terwee CB et al (2009) Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Quality of Life Research* 18:8;1115–1123
- The George Institute for Global Health (2014) PEDro Physiotherapy Evidence Database. <http://www.pedro.org.au/> (accessed 07-04-2015)
- UK Department of Health. Reference costs. <https://www.gov.uk/government/publications/nhs-reference-costs-2012-to-2013>
- University of York Library (2015) Find health economics resources. <http://subjectguides.york.ac.uk/healtheconomics> (accessed 26-06-15)
- Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MMG, Sterne JAC, Bossuyt PMM, the QUADAS-2 Group. (2011). QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 155(8):529–36
- Wilczynski NL, Haynes RB, Lavis JN, Ramkissoonsingh R, Arnold-Oatley AE, HSR Hedges team (2004) Optimal search strategies for detecting health services research studies in MEDLINE. *Can Med Assoc J* 171(10):1179–85

# Chapter 3

## Building a Decision Tree

### Cost Effectiveness Model

**Abstract** Decision analytical modelling as a vehicle for cost effectiveness analyses may use various modelling approaches including decision trees and Markov models. Determining when to use a particular modelling approach and choice of model will depend on a number of different factors. For example, decision trees are most useful when health events happen close together and don't repeat; when health events happen quickly or not at all; and when uncertainty over the effects of treatment is resolved quickly. This chapter guides you through choice of model with focus lying on how to develop a decision tree to assess cost effectiveness.

### 3.1 Introduction

Decision analytical modelling is often described as a *vehicle* for cost effectiveness analysis (CEA) (Saramago et al. 2012). As highlighted in the previous chapter, whilst CEA may be undertaken within the context of a single randomised controlled trial (RCT), use of single trial data alone limits analyses to the population within that trial, the interventions assessed and the time over which the trial participants are followed up (NICE 2013). Decision analytical modelling allows us to tailor our analyses to the question we are trying to answer, facilitated by use of multiple sources of data to answer questions relating to resource allocation that are typically beyond the scope of a single RCT.

This chapter introduces decision trees in cost effectiveness models and is structured as follows: In Sect. 3.2 we define decision modelling and discuss when decision modelling in CEA is appropriate and what factors influence choice of model (decision tree or Markov model). Section 3.3 then guides you through how to develop a decision tree to assess cost effectiveness. Using an example, we show you how to construct and populate a model and how to interpret the results. Section 3.4 discusses potential costs and outcomes that you might want to think

about including and the level of complexity of your model. Having guided you through constructing a simple decision tree, the exercise in Sect. 3.5 offers opportunity for you to build one yourself. The chapter finishes on a summary of the key points in Sect. 3.6.

## 3.2 What Is a Decision Model?

Within health and the health care sector, cost effectiveness analyses, and economic evaluations more generally, are undertaken to inform resource allocation decisions. CEA is used to inform choices about what health care to provide, for whom and when. It allows decision makers to consider whether the provision of a health care treatment, programme or technology represents value for money. Modelling in cost effectiveness provides a common framework by which to systematically assess costs and outcomes, to represent uncertainty and to compare alternative courses of action. More formally, decision modelling can be defined as *a systematic approach to decision making under conditions of uncertainty, in which the probability of each possible event, along with the consequences of those events, is explicitly stated* (Kielhorn and von der Schulenburg 2000).

Thus the modelling framework is *systematic* because outcomes are incorporated within a common framework; it takes account of *uncertainty* because we do not know exactly what will happen, and it includes the *consequences* of alternative actions (incremental analysis). The *decision* relates to resource allocation decisions, comparing different courses of action against each other.

The choice to use modelling within CEA is driven by a number of factors. The box below illustrates when decision-based modelling is appropriate to use (Box 3.1).

We can see that modelling is often used in the absence of a relevant or appropriate RCT. This may be simply that no RCT has been conducted that addresses the

### Box 3.1 When Decision Modelling is Appropriate to Use (NICE 2013)

- All the relevant evidence is not contained in a single trial
- Patients participating in trials do not match the typical patients likely to use the technology
- Intermediate outcome measures are used in trials rather than effect on health-related quality of life and survival
- Relevant comparators have not been used, or trials do not include evidence on relevant subgroups
- Clinical trial design includes crossover (treatment switching) that would not occur in clinical practice
- Costs and benefits of the technologies extend beyond the trial follow-up period

question posed, but as Box 3.1 suggests, there may be other factors at play. For example, RCTs are designed to answer a specified question that may not reflect *real life* in as much as the strict participant criteria for the trial are likely to mean that the patient population is narrower than would be the case in typical practice.<sup>1</sup> RCTs are necessarily time limited so potentially important longer-term consequences may not be recorded. Additionally, the health care intervention of interest may not be compared against a relevant or credible alternative in the current context. If one or more of the statements included in Box 3.1 are true, then it is likely that a decision model is the appropriate form of CEA.

Whilst the RCT might not reflect typical practice, one of the dichotomies associated with decision modelling is that whilst the decision model itself should mirror reality, there is a tension between the complexity of *real life* and the need for any model to be simple enough to be useful, *doable* in terms of construction and computational needs and understandable to the person using it. We will come back to this later in the chapter.

Having decided whether a model is appropriate for the CEA, what criteria do we use to choose the type of model? Whilst there are a number of different types of decision analytical models including, for example, discrete event simulation modelling, the focus of this book is decision trees and Markov models. Choice of decision tree or Markov model is relatively clear cut. Decision trees are most useful when health events happen close together and don't repeat; when health events happen quickly or not at all; and when the effects of treatment are over quickly. Conversely, Markov models are most useful when health events repeat over time or have longer-term health effects; when there are a large number of potential health effects over time; and when the risk of different health events does not depend on the patient's prior history. Given these criteria, in order to decide what type of model is most appropriate, it is essential to understand the natural history of the health condition that will be the subject of the decision model and the care pathway. This means asking do health events happen over time? Do the risks of events change over time? Is the intervention equally effective over time? Does an initial health change persist once treatment stops? And, does the risk of different health events depend on a patient's history? Sources of information of the natural history and care pathway include evidence from existing research, clinical guidelines and expert opinion. Chapter 2 discusses both the sources and how to search for this evidence in more detail.

Once you are clear about the natural history and care pathway of the condition of interest and you have decided on the most appropriate type of model, you can begin the task of building the model.

---

<sup>1</sup> As a broader point, it is useful to distinguish between efficacy and effectiveness (although they do exist on a continuum). The efficacy of an intervention is the extent to which an intervention is effective when studied under controlled research conditions. Effectiveness is the extent to which an intervention produces an overall health benefit in routine clinical practice (NICE 2013).

### 3.3 Key Elements of a Decision Tree

The first step in building a decision tree, and in fact any decision model, is formulating the decision problem. You will need to think carefully about the question you want to ask and how to frame it. The importance of this step cannot be overstated as it will inform all subsequent steps of building the decision tree – and of course the results and the interpretation of those results. The decision problem should involve at least two options and at least one outcome upon which to base a recommendation. The question should also normally indicate the group that is affected.

Consider the value of two different types of hip replacement surgery for young people suffering from arthritis: total hip arthroplasty where the whole joint is replaced and resurfacing arthroplasty where part of the joint is replaced, making a smooth surface to allow more normal hip function. We might want to know which type of surgery leads to the best recovery. In this case we might formulate our decision problem, our question, in the following way:

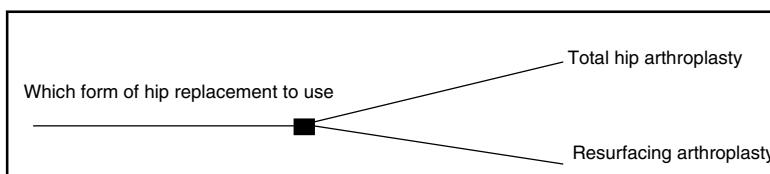
*Which method of operating on arthritic hips in young patients, total hip arthroplasty or resurfacing arthroplasty provides the likelihood of good recovery?*

In this case the two options are total hip arthroplasty and resurfacing arthroplasty, the outcome is the likelihood of good recovery and the group affected is young patients.

We can now begin the process of building our decision tree. Decision trees are read from left to right and begin at a single point on the left-hand side. We begin with a decision or choice, and this is shown in the decision tree by a decision node represented by ■. As the name suggests, the decision node indicates where a decision is made, and the lines or branches of the decision tree emanating from this node show the options at this point. Figure 3.1 shows the beginning of our decision tree; the branches to the right of the decision node show the options of total hip arthroplasty and resurfacing arthroplasty in line with our decision problem. Whilst there are only two options in this example, for other questions it may be appropriate to include more.

Having identified the options, we can build on this using a chance node represented by ●. The chance node defines a risk and indicates what will happen as a result of it. In this example, we have identified the two options for surgery. The two possible outcomes as a result of that surgery are that the patient may die whilst having the operation (perioperative death) or survive. These can be added to the decision tree. This is illustrated in Fig. 3.2.

Now we have added survival and perioperative death; we next need to think about the chance or probability of each occurring; these are known as the transition probabilities – the probability of transitioning or moving into the survival branch or



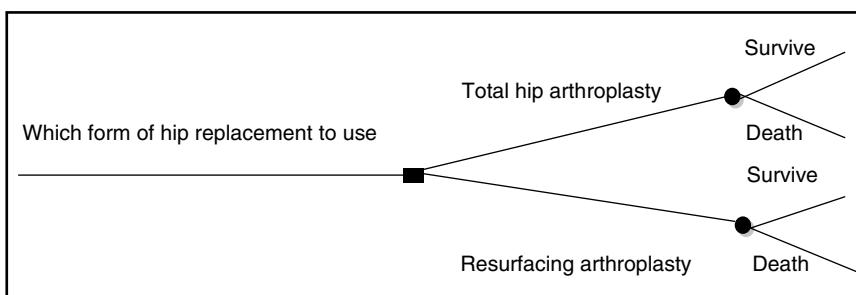
**Fig. 3.1** Decision trees: starting out

the perioperative death branch. The probabilities in the branches from each chance node must add to one. For this example we have assumed the probabilities are as follows: For patients undergoing a total hip arthroplasty, the probability of surviving the surgery is 99 %, and thus the probability of death would be 1 %. For patients undergoing resurfacing arthroplasty, the probability of survival is 98.5 %, and the probability of death is 1.5 %. These are shown in Fig. 3.3.

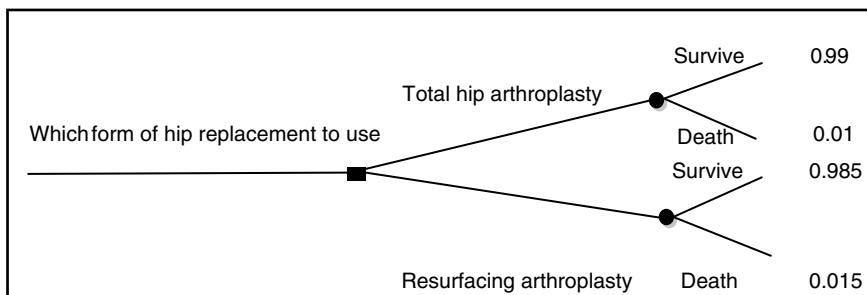
Whilst Fig. 3.3 shows us the probability of survival and death in surgery for each type of surgical procedure, the structure is incomplete – it doesn't yet answer our question. Remember we want to know which procedure *provides the likelihood of good recovery*. At present we know the likelihood of survival, but we don't yet know the likelihood of a *good recovery* if the patient survives surgery. We can now add this. For those patients who have survived, we have another chance node – in this case it indicates the probability that the patient will have good function or poor function having survived surgery. In this case we assume function is mobility, and that good function means no impairment to mobility whilst bad function is impaired mobility, for example, difficulty climbing stairs.

We can also add terminal nodes to our diagram. The terminal nodes are used where there is no more risk and are represented by ▲. Terminal nodes indicate that we have observed the outcomes we are looking for, including but not limited to the case when someone has died. In our example, we have added good function and bad function, so we have included the outcome we need to answer the question in terms of the likelihood of a good recovery.

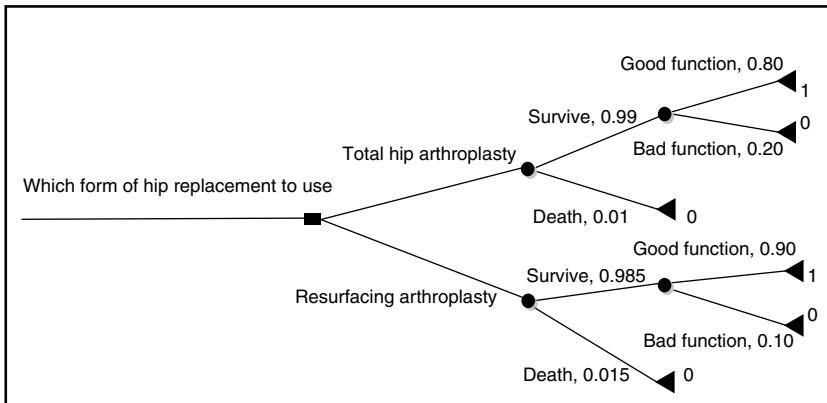
At each of the terminal nodes, payoffs are defined. These might include the costs of health care and/or the value of the health state. For example, the health care costs



**Fig. 3.2** Decision trees: adding chance nodes



**Fig. 3.3** Decision trees: adding probabilities



**Fig. 3.4** Decision trees: adding terminal nodes and payoffs

might include the costs of the surgery, medications and rehabilitation. The health outcome might be a measure of quality of life or quality adjusted life years (QALYs). There is normally more than one payoff (costs and health outcome). However, for this simple example we have assumed single payoffs. The payoff for a good function is 1, and the payoff for poor function or perioperative death is 0. Figure 3.4 shows the addition of a chance node with good and poor function; terminal nodes have been added to the perioperative death and to good and poor function. In addition payoffs are added to each of the terminal nodes.

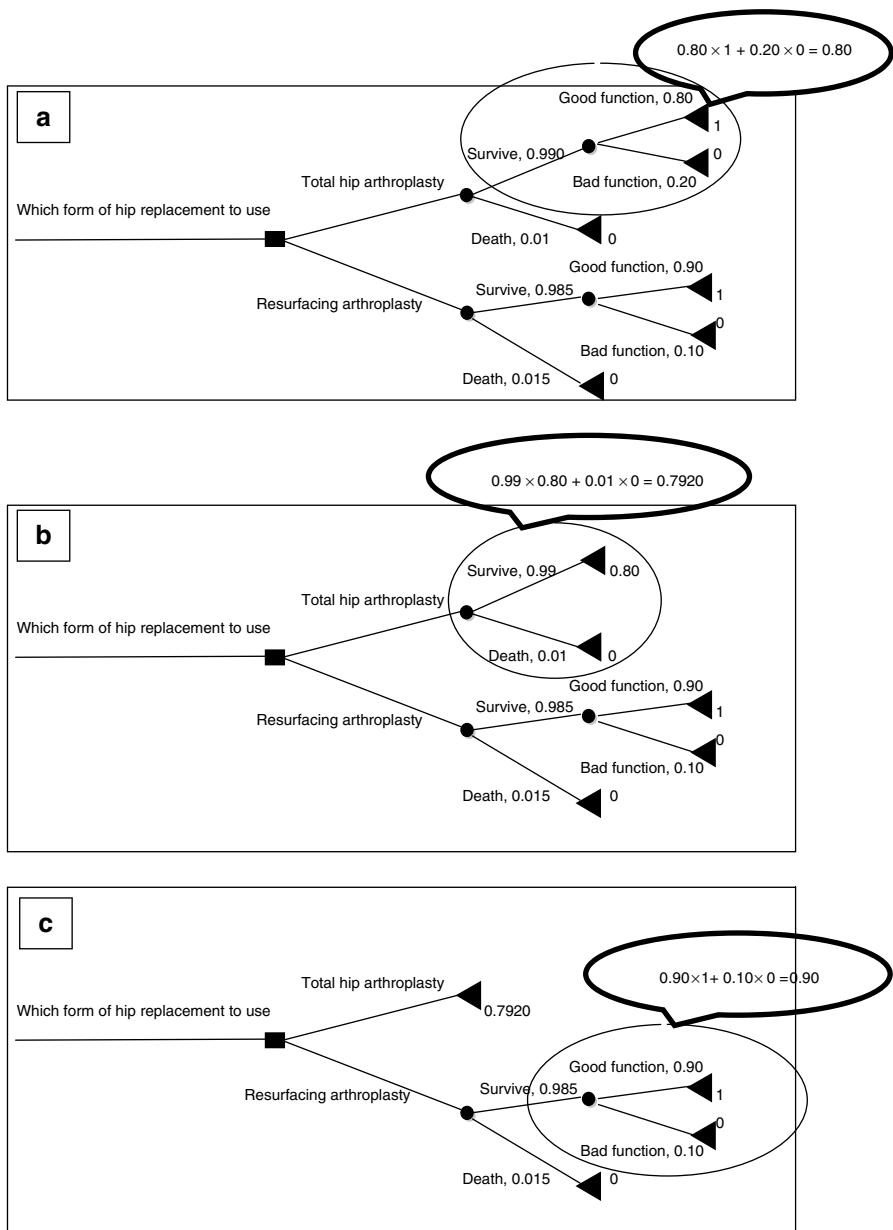
Now we have the structure of our decision tree, we have the transition probabilities and the payoffs and are ready to calculate the expected value of the likelihood of a good recovery for total hip arthroplasty and for resurfacing arthroplasty. We estimate the expected payoff of each option, that is, the payoff values at each terminal node weighted by the unconditional probability of reaching that node.

In our example, we can follow the pathways from the original decision node (total hip arthroplasty or resurfacing arthroplasty) through to each of the terminal nodes. We will first consider the total hip arthroplasty branch of the tree. For patients undergoing total hip arthroplasty, there are three distinct pathways: good function, poor function and perioperative death. We can calculate the likelihood of the payoff for each of these pathways in the following way:

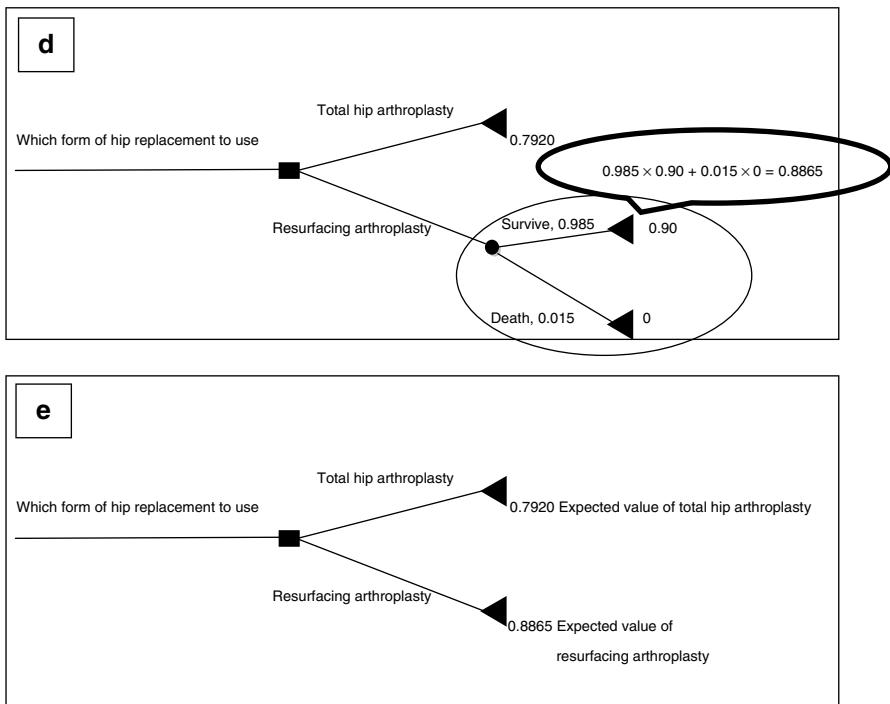
$$\begin{aligned} &\text{Likelihood of payoff for good function} \\ &= \Pr(\text{survival}) \times \Pr(\text{good function}) \times \text{payoff} \\ &= 0.99 \times 0.80 \times 1 = 0.7920 \end{aligned}$$

$$\begin{aligned} &\text{Likelihood of payoff for bad function} \\ &= \Pr(\text{survival}) \times \Pr(\text{bad function}) \times \text{payoff} \\ &= 0.99 \times 0.20 \times 0.00 = 0.00 \end{aligned}$$

$$\begin{aligned} &\text{Likelihood of payoff for perioperative death} \\ &= \Pr(\text{perioperative death}) \times \text{payoff} \\ &= 0.01 \times 0.00 = 0.00 \end{aligned}$$



**Fig. 3.5** Decision trees: expected value. (a) Expected value for patients surviving total hip arthroplasty. (b) Expected value of total hip arthroplasty. (c) Expected value for patients surviving resurfacing arthroplasty. (d) Expected value of resurfacing arthroplasty. (e) The expected value of a good recovery

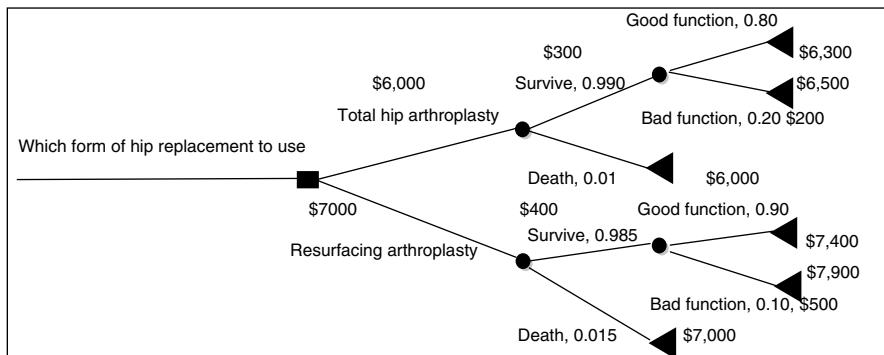


**Fig. 3.5** (continued)

The *expected value* of total hip arthroplasty is the likelihood of the payoff of the three pathways added together, which is  $0.7920$  ( $0.7920+0+0$ ). We can now calculate the expected value for patients undergoing resurfacing arthroplasty in the same way. In this case the expected value is  $((0.985 \times 0.9 \times 1) + (0.985 \times 0.10 \times 0) + (0.015 \times 0)) = 0.8865$ . The calculations are shown step by step in the diagrams above (Fig. 3.5).

As we have calculated the expected value for each of the surgery options, we can now answer our question: which method of operating on arthritic hips in young patients, total hip arthroplasty or resurfacing arthroplasty provides the greater likelihood of good recovery? Looking at the expected value from each surgical option, we can see that resurfacing arthroplasty has a higher likelihood of a good recovery. The expected value from resurfacing arthroplasty is 0.8865 compared with 0.7920 for total hip arthroplasty. However, you might have noticed that in this example that whilst the calculation shows overall a better recovery, defined as a higher probability of good function for those who survive compared to survivors of total hip arthroplasty, the probability of perioperative death is higher for the resurfacing surgery.

In general, economic models assume the decision maker is risk neutral, that they are indifferent to options with the same expected value even if one is more uncertain or riskier than another. In our example there is clearly a higher risk of death associated with the resurfacing surgery. Defining the decision problem in terms of the likelihood of good function has led us to ignore the difference between being alive with a bad function and being dead. This shows us not only how important formulating the ques-



**Fig. 3.6** Decision trees: expected costs

tion is, but also the importance of how we interpret the findings. In practice decision makers take into account a number of factors when deciding on resource allocation and as such must decide how much weight they place on the results of the economic model.

### 3.4 Costs, Benefits and Complexity

The example in the previous section used single values. However, for an economic evaluation using a decision tree, the payoffs will need to include both costs and benefits. In line with any economic evaluation, you must consider the perspective the evaluation is taking. This will be informed by the question being asked and the purpose of the analysis. The costs might include costs to the health care system, costs to the economy in general or even costs tailored to the specifications of a particular decision maker.<sup>2</sup> Let's look at each of these in turn. Costs to the health system might include the costs to secondary, tertiary or primary health care and may in some cases be extended to include costs of personal social care provision. If a wider societal perspective is taken, this might widen the scope to include not only health and personal social care but also productivity arising from days away from work, the costs of informal care and out-of-pocket expenses accruing to the patient and their family. Alternatively the analysis may focus on the impact of a particular intervention in a specified area, for example, hospital admissions.

Figure 3.6, above, shows the same process, but this time we have attached the costs associated with each option. We assume a cost of \$6,000 for the initial total hip arthroplasty operation and a cost of \$7,000 for the initial resurfacing arthroplasty operation. If the patient dies perioperatively, no further costs are accrued. For those patients who survive with good function, there will be further outpatient visits attracting a cost of \$300 for those who have had the total hip arthroplasty and \$400

<sup>2</sup> A number of excellent texts are available that describe in more detail the principles of economic evaluation; these range from generic health economics texts such as Morris et al. (2012) to those whose focus lies entirely on economic evaluation such as Drummond et al. (2005).

for those who had the resurfacing arthroplasty. For those with bad function, there will be the same outpatient visits attracting the same cost, but additionally there will be subsequent inpatient days attracting a cost of \$200 for those who had total hip arthroplasty and \$500 for those who had the resurfacing arthroplasty operation. The costs are entered into our decision tree. Thus we can see the costs in the total hip arthroplasty arm are \$6,300, \$6,500 and \$6,000; in the resurfacing arm they are \$7,400, \$7,900 and \$7,000.

Using the same formula as before, we can calculate the expected costs.

For the total hip arthroplasty arm:

$$\begin{aligned} &\text{Likelihood of payoff for good function} \\ &= \text{Pr(survival)} \times \text{Pr(good function)} \times \text{payoff} \\ &= 0.99 \times 0.80 \times \$6,300 = \$4,989.60. \end{aligned}$$

$$\begin{aligned} &\text{Likelihood of payoff for bad function} \\ &= \text{Pr(survival)} \times \text{Pr(bad function)} \times \text{payoff} \\ &= 0.99 \times 0.20 \times \$6,500 = \$1,287. \end{aligned}$$

$$\begin{aligned} &\text{Likelihood of payoff for perioperative death} \\ &= \text{Pr(perioperative death)} \times \text{payoff} \\ &= 0.01 \times \$6,000 = \$60. \end{aligned}$$

The expected cost of total hip arthroplasty is the likelihood of the payoff of the three pathways added together, which is  $\$4,989.60 + \$1,287 + \$60 = \$6,336.60$ .

For the resurfacing hip arthroplasty arm:

$$\begin{aligned} &\text{Likelihood of payoff for good function} \\ &= \text{Pr(survival)} \times \text{Pr(good function)} \times \text{payoff} \\ &= 0.99 \times 0.80 \times \$7,400 = \$5,860.80 \end{aligned}$$

$$\begin{aligned} &\text{Likelihood of payoff for bad function} \\ &= \text{Pr(survival)} \times \text{Pr(bad function)} \times \text{payoff} \\ &= 0.99 \times 0.20 \times \$7,900 = \$1,564.20 \end{aligned}$$

$$\begin{aligned} &\text{Likelihood of payoff for perioperative death} \\ &= \text{Pr(perioperative death)} \times \text{payoff} \\ &= 0.01 \times \$7,000 = \$70 \end{aligned}$$

In this case the expected cost is  $\$5,860.80 + \$1,564.20 + \$70 = \$7,495$ . Thus the expected cost of total hip arthroplasty is lower than for resurfacing arthroplasty.

We can now calculate the ICER:

$$\begin{aligned} \text{ICER} &= \frac{C_2 - C_1}{E_2 - E_1} = \frac{\$7,495 - \$6,336.60}{0.8865 - 0.7920} = \frac{\$1,158.40}{0.0945} \\ &= \$12,258 \text{ per Good Functioning Hip} \end{aligned}$$

In this case, for the purpose of simplicity, we assigned arbitrary payoffs in respect of the benefits of surgery (1 for good function and 0 for either bad function or death), but when designing a decision tree, you will need to think much more about the outcome you will use. This will be informed by the question you are asking. In our hip surgery

example, we saw that whilst survival was represented within the model, focus lay on the likelihood of a *good* recovery. As such whilst perioperative death was important and we needed to identify the probability of death associated with each type of surgery, we also needed to include an outcome that gave a measure of recovery (in this case we used function). Within any decision tree, you might choose outcomes that are mortality based (e.g. life years, life expectancy, mortality), a clinical outcome (e.g. cases of cancer survival at 5 years), intermediate or proxy outcomes (e.g. cases detected or biomarkers), a quality of life outcome (e.g. self-assessed health) or a composite outcome such as QALYs. Further details of the type of evidence or data you might use to populate your model and how you search for it are covered in Chap. 2.

We said earlier that the importance of formulating your research question cannot be overstated as it will inform all subsequent steps of building the decision tree. The question chosen together with factors such as the natural history of the disease will feed into how complex the model is. There is a tension between the complexity of real life and the need for any model to be simple enough to be useful and understandable to the person using it. How then do you decide whether your model has become too complex and whether this is a problem? Well, you should think about both practical and conceptual implications. On a practical level, you will need to take into account the resources you have available to build the model and populate it and whether the model has become too big and unwieldy (your decision tree has become too *bushy*!) and whether the level of detail means that you can no longer get information to fill the tree. For example, are the subgroups you are looking at so small that no data is available? Conceptually think about how complex your model needs to be in order to answer the question posed. Consider whether the level of detail in the model means that decision makers would struggle to distinguish between the options. If when building your model you have started to worry about *when* things happen or if the timeframe extends beyond one year, then you should think about whether a decision tree is the right type of model as decision trees are not designed to deal with complex changes over time. When we considered the criteria for an appropriate model type, we stated that decision trees are most useful when health events happen close together and don't repeat; when health events happen quickly or not at all; and when the effects of treatment are over quickly. If you decide that your proposed model structure is too complex, think about how you might simplify it or indeed if the decision problem itself can be simplified. This is always a difficult call, but consider paring back your model or the question to reduce the complexity and the value of doing so. If calculating your outcomes means that you need to know when particular events occur, then you should strongly consider whether a Markov model might be a more appropriate form of model.

### 3.5 Exercise Building a Decision Tree

The following exercises provide the opportunity for you to build a decision tree. The first exercise uses pen and paper, whilst Exercises 3.2 and 3.3 use Excel. Excel worksheets for Exercises 3.2 and 3.3 can be downloaded from <http://medhealth.leeds.ac.uk/costeffectivenessmodelling>. All three use information from a Canadian study provided in a paper from Evans et al. (1997). In the paper the decision model

is based around a brief description of what might happen with a person who suffers from recurrent migraines, on one of the days that they suffer from a migraine. Two options are compared – giving the patient an oral dose of sumatriptan or a combined oral dose of caffeine and ergotamine.

The basic ‘story’ is the same for both treatments:

- After treatment is given, the patient can either get relief (the patient becomes free from pain after 2 h) or no relief at all.
- In many cases where the patient gets some relief, he/she will not suffer any further.
- In other cases, he/she may get some relief but have another attack later and will require retreatment, which is assumed to be effective.
- If the patient does not get relief from the initial treatment, he/she will either stay at home and suffer from the migraine or go to the hospital. If she goes to the hospital’s emergency room for treatment, she will either get relief from treatment or be admitted to the hospital where this does not happen.

### **Exercise 3.1 Drawing a Decision Tree**

In Exercise 3.1 your task is to draw the skeleton of a decision tree corresponding to the description above. Once you have drawn the decision tree, you can add the transition probabilities based on the following information:

- The probability of conversion of a moderate or severe headache to a mild headache or no headache at all within 2 h with (1) sumatriptan is 55.8 % and (2) caffeine/ergotamine is 37.9 %.
- The probability of having another attack within 48 h after relief with (1) sumatriptan is 40.6 % and (2) with caffeine/ergotamine is 29.7 %.
- If the medication does not relieve the migraine, the probability of going to the hospital emergency room is 8 %.
- If the treatment received in the emergency room does not relieve the migraine, there is a 2 % chance of hospitalisation.

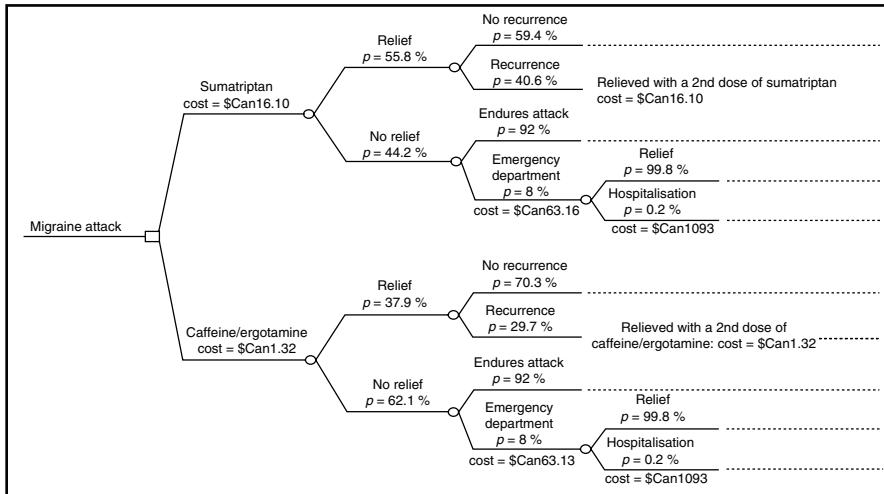
Now add the probabilities to your diagram. Once this is complete add the following costs to the diagram. The individual costs can be written against the branch of the model in which they occur.

- Sumatriptan Can\$16.10
- Caffeine/ergotamine Can\$1.32
- Emergency room Can\$63.13
- Hospitalisation Can\$1,093

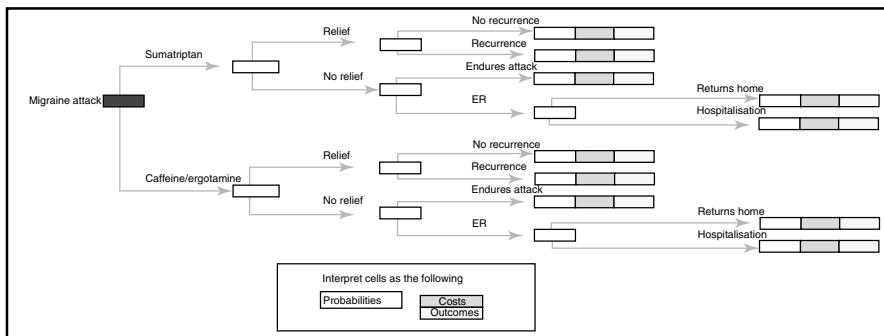
N.B. For those people who have a recurrence, we assume another dose of the same medication is taken. Once you have included the costs, you can check how you did by looking at the decision tree in Fig. 3.7.

### **Exercise 3.2 Building a Decision Tree in Excel(1)**

As described in Exercise 3.1, the Evans et al. model is based around what might happen with a person that suffers from recurrent migraines, on one of the days that they suffer from a migraine. To recap, two options are compared, giving the patient an oral dose of sumatriptan or a combined oral dose of caffeine and ergotamine. The economic model aims to find an incremental cost effectiveness ratio. If you have left



**Fig. 3.7** Decision tree showing sumatriptan or a combined oral dose of caffeine and ergotamine for migraine relief



**Fig. 3.8** Migraine decision tree in Excel

time between exercises 3.1 and 3.2, it might be useful to revisit Exercise 3.1 to (re) familiarise yourself with the description of the care pathway. The decision tree, including the probabilities and costs included in the exercise, is shown below (Fig. 3.7).

Whilst it is useful to draw a decision tree using pen and paper, for the most part, this type of model is usually built, and the analysis runs, in software packages such as Excel. In this exercise we talk you through building the decision tree model in Excel. The first step is to draw the decision tree. Figure 3.8 sets the problem out as a decision tree mirroring the structure of the tree above. The unshaded cells with a black border record the proportion of patients in each node. The set of three cells in bold capture the proportion of patients in each node in the first cell, the outcomes associated with patients on each pathway in the middle cell and the costs in the final cell.

Open a second worksheet in the same Excel workbook that you have constructed the decision tree (Fig. 3.8). On this new sheet, set up tables corresponding to the cells in the first worksheet. These are shown in Figs. 3.9 and 3.10.

	Sumatriptan			Caffeine/ergotamine		
	Prob	Cost	Outcomes	Prob	Cost	Outcomes
All cases		0.000			0.000	
Initial relief	■	0.000		■	0.000	
No recurrence	0	0	0	0	0	0
Recurrence	0	0	0	0	0	0
No relief	■	0.000		■	0.000	
Endures attack	0	0	0	0	0	0
ER	■	0.000		■	0.000	
Returns home	0	0	0	0	0	0
Hospitalisation	0	0	0	0	0	0

**Fig. 3.9** Probabilities, costs and outcomes in table form in Excel

	Sumatriptan			Caffeine/ergotamine		
	Prob	Cost	Outcomes	Prob	Cost	Outcomes
No recurrence	0	0	0	0	0	0
Recurrence	0	0	0	0	0	0
Endures attack	0	0	0	0	0	0
Returns home	0	0	0	0	0	0
Hospitalisation	0	0	0	0	0	0
Expected values						

**Fig. 3.10** Probabilities, costs and outcomes in table form in Excel

We will now focus on the first worksheet – the decision tree model. If we choose the sumatriptan option, there is a 100 % chance of beginning at the risk node in that arm (circle). Enter 1 into the sumatriptan starting cell. You can now use the probabilities in the table to place values into the subsequent nodes. Enter probabilities for relief/no relief based on the information we were given. For those who have relief, they will have either no recurrence or a recurrence, and we can complete these cells. Now do the same for the possibilities following no relief, filling in values in the appropriate cells. Fill in the rest of the sumatriptan probabilities, and do the same for the caffeine/ergotamine side of the tree.

We have now filled in the chances of reaching each terminal node – the places that the tree ‘stops’. We now need to identify the costs attached to each terminal node too. The tree also contains information about the costs of treatment, and these relate to (a) the cost of sumatriptan (plus retreatment), (b) the cost of caffeine/ergotamine (plus retreatment), (c) the cost of an emergency department visit and (d) the cost of hospitalisation following an emergency department visit.

		Sumatriptan			Caffeine/ergotamine		
		Prob	Cost	Outcomes	Prob	Cost	Outcomes
All cases		1.000			1.000		
Initial relief		0.558			0.379		
No recurrence		0.331452	\$ 16.10	1.00	0.266437	\$ 1.32	1.00
Recurrence		0.226548	\$ 32.20	1.00	0.112563	\$ 2.64	1.00
No relief		0.442			0.621		
Endures attack		0.40664	\$ 16.10	-	0.57132	\$ 1.32	-
ER		0.035			0.050		
Relief		0.035289	\$ 79.26	-	0.049581	\$ 64.48	-
Hospitalisation		7.07E-05	\$ 1,172.23	-	9.94E-05	\$ 1,157.45	-
Expected values			\$ 22.06	0.56		\$ 4.71	0.38

		Sumatriptan			Caffeine/ergotamine		
		Prob	Cost	Outcomes	Prob	Cost	Outcomes
No recurrence		0.331452	\$ 16.10	1.00	0.266437	\$ 1.32	1.00
Recurrence		0.226548	\$ 32.20	1.00	0.112563	\$ 2.64	1.00
Endures attack		0.40664	\$ 16.10	-	0.57132	\$ 1.32	-
Relief		0.035289	\$ 79.26	-	0.049581	\$ 64.48	-
Hospitalisation		7.07E-05	\$ 1,172.23	-	9.94E-05	\$ 1,157.45	-
Expected values			\$ 22.06	0.56		\$ 4.71	0.38

**Fig. 3.11** Completed tables

For the first terminal node, identify the costs incurred to reach the sumatriptan ‘no recurrence’ node by following the tree from ‘sumatriptan’ to ‘no recurrence’. Enter this cost into cost cell for the sumatriptan no recurrence pathway. Now, do the same for the other terminal nodes, both sumatriptan and caffeine/ergotamine.

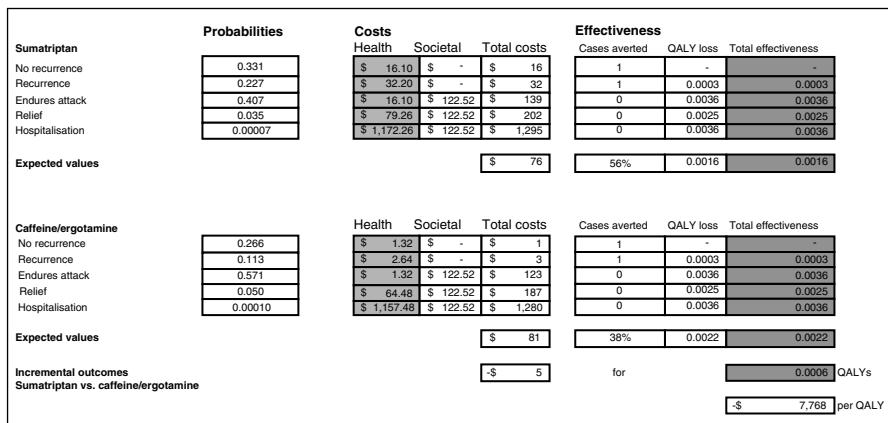
Now for outcomes, Evans uses two different outcomes – attacks averted and quality adjusted life years. For simplicity, we’ll use the attacks averted. Relief after an ER attendance was not treated as relief by Evans. For each terminal node, decide whether or not the attack was averted. Place a 1 if the attack was averted and a 0 if the attack was not averted.

The figures entered on the decision tree can now be entered into the tables in the second worksheet. Alternatively you can link the two so that the figures are automatically transferred. Whilst providing the same information, the tables show the figures in a much more compact format. This compact form is more typical when building a model – however, it is helpful to see how you get there.

It is now time to calculate some expected outcomes. As in Fig. 3.10 you will now have the expected proportion of patients in each terminal node, the costs and the cases averted (outcomes). The next step is to calculate expected values for these states in the usual way. What can you conclude about the costs and benefits of sumatriptan vs. caffeine/ergotamine? Check your answers against the completed tables below (Fig. 3.11).

**Table 3.1** Quality of life outcomes

	Quality of life	Standard deviation
Attack averted with initial treatment	1.0	0
Attack averted with retreatment	0.9	0.01
Attack endured after failed treatment	-0.3	0.10
Emergency room visit after failed treatment	0.1	0.10
Hospitalisation after failed treatment	-0.3	0.10

**Fig. 3.12** Decision tree using QALYs**Exercise 3.3 Building a Decision Tree in Excel(2)**

The Evans model used both ‘attack averted’ and QALYs as outcome measures. The QALY-based outcomes assume that each attack lasts 24 h and that quality of life is as follows (Table 3.1):

In this exercise, run the model again using QALYs instead of cases averted. Use the values in Table 3.1 to replace the cases averted (as the outcome measure). You can calculate the QALY loss from a migraine as  $(1-\text{quality of life})/365$ . The Evans paper suggested that sumatriptan produced around 0.0006027 QALYs more. Can you replicate this result? The answer is shown in the Fig. 3.12 above.

## 3.6 Summary

- Decision analytical modelling is a vehicle for CEA and may use various modelling approaches.
- Decision models are most appropriate when all the relevant evidence is not contained in a single trial; patients participating in trials do not match the typical

patients likely to use the technology; intermediate outcome measures are used rather than effect on health-related quality of life and survival; relevant comparators have not been used or trials do not include evidence on relevant subgroups; clinical trial design includes crossover (treatment switching) that would not occur in clinical practice; and/or when costs and benefits of the technologies extend beyond the trial follow-up period (National Institute for Health and Care Excellence 2013).

- Decision trees are most useful when health events happen close together and don't repeat; when health events happen quickly or not at all; and when uncertainty over the effects of treatment are resolved quickly.
- The decision problem should involve at least two options and at least one outcome upon which to base a recommendation. The question should also normally indicate the group that is affected.
- Decision trees are read from left to right and begin at a single point on the left-hand side. We begin with a decision or choice, and this is shown in the decision tree by a decision node. This indicates where a decision is made and the lines or branches of the decision tree emanating from this node show the options at this point.
- At the end of these branches is a chance node which defines a risk and indicates what will happen as a result of it. The probabilities in the branches from each chance node must add to one.
- Terminal nodes are used where there is no more risk. They occur either because we have observed the outcomes we are looking for or because someone has died. At each of the terminal nodes, payoffs are defined. The payoffs will need to include costs and benefits.
- There is a tension between the complexity of real life and the need for any model to be simple enough to be useful and understandable to the person using it.

## References

- Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ (2005) Methods for the economic evaluation of health care programmes, 3rd edn. OUP, Oxford
- Evans KW, Boan JA, Evans JL, Shuaib A (1997) Economic evaluation of oral sumatriptan compared with oral caffeine/ergotamine for migraine. *Pharmacoeconomics* 12(5):565–577
- Kielhorn A, von der Schulenburg JMG (2000) The health economics handbook. Adis International, Chester
- Morris S, Devlin N, Parkin D, Spencer A (2012) Economic analysis in health care. Wiley, Chichester
- National Institute for Health and Care Excellence (2013) Guide to the methods of technology appraisal 2013. <http://publications.nice.org.uk/pmg9>
- Saramago P, Manca A, Sutton AJ (2012) Deriving input parameters for cost-effectiveness modeling: taxonomy of data types and approaches to their statistical synthesis. *Value Health* 15:639–649. doi:[10.1016/j.jval.2012.02.009](https://doi.org/10.1016/j.jval.2012.02.009)

# **Chapter 4**

## **Uncertainty, Probabilistic Analysis and Outputs from Cost Effectiveness Analyses**

**Abstract** An economic evaluation provides a summary of the current evidence for different actions in response to a decision problem, and aims to inform decision makers about what impacts these might have in terms of costs and benefits. As many of these costs and benefits occur in the future, and the future is inherently uncertain, it is important that economic evaluations explore the uncertainties in any answer that can be produced from current evidence. This chapter focuses on introducing the different sources of uncertainty in economic evaluation and the more straightforward mechanisms for their characterisations and analysis, along with a critical understanding of the limitations of these methods.

### **4.1 Introduction**

The ultimate purpose of any economic evaluation is to inform decision-making; it is inherently concerned with estimating the expected future costs and outcomes of alternative courses of action; and because it is concerned with the future, uncertainty is a central concern for the analyst. It is impossible to be completely certain about the future. The question that decision makers look to analysts to help them answer is: ‘Is the nature or the magnitude of the uncertainty such that there is a substantial risk that I will make the wrong decision using the available information?’ This risk of making the wrong decision is often referred to as ‘decision uncertainty’.

It is important to note that uncertainty that if resolved could not change the decision and does not contribute to decision uncertainty. For example, where we are not uncertain about which alternative is best, but are only uncertain about *how much* better a clearly superior option is, there is no decision uncertainty. In

the context of reimbursement decision-making in health care, decision uncertainty can be thought of as the probability that the decision maker will refuse to pay for a therapy that is good value or agree to pay for a therapy that is not good value.

For the analyst to help the decision maker answer this question, it is necessary to understand the different sources of uncertainty that can impact upon the results of an economic evaluation, the methods that are available to characterise each type of uncertainty and the mechanisms for analysing the decision uncertainty that each produces. Over the course of the book, we will address each of these issues in some detail. However, the focus of this chapter is to introduce the different sources of uncertainty in economic evaluations and the more straightforward mechanisms for their characterisation and analysis, along with a critical understanding of the limitations of these methods. The chapter is structured as follows: Sect. 4.2 describes the sources of uncertainty; Sect. 4.3 discusses the methods by which you as an analyst can address uncertainty; Sect. 4.4 explores some of the problems associated with the use of incremental cost effectiveness ratios (ICERs) expanding on the cost effectiveness planes and Net Benefit calculation described in Chap. 1. The final section contains a summary of the key points.

## 4.2 Sources of Uncertainty in Cost Effectiveness Models

Five distinct sources of uncertainty need to be considered in cost effectiveness modelling: sampling variation, extrapolation, generalisability, model structure and methodological uncertainty.

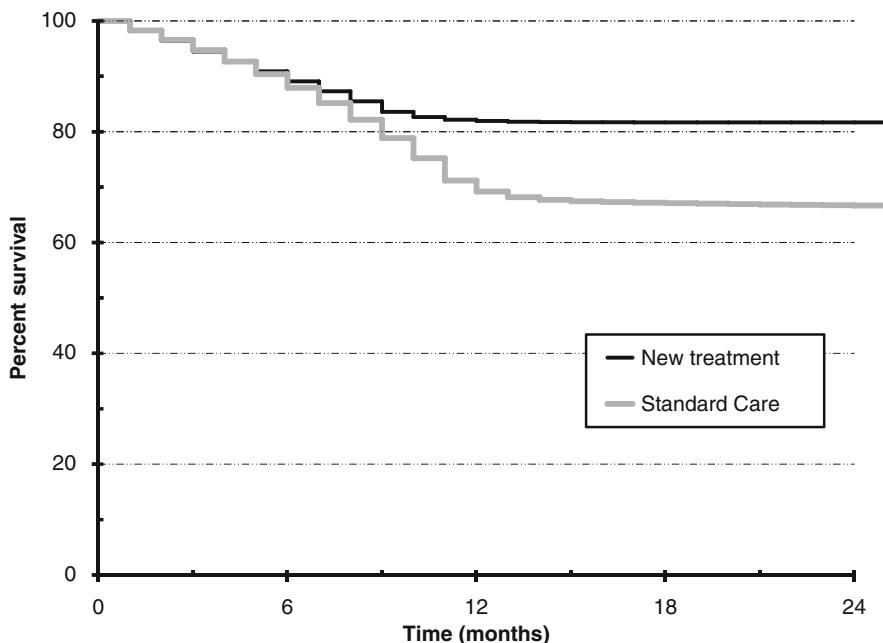
### 4.2.1 Sampling Variation

The vast majority of the information synthesis in a cost effectiveness model will be obtained from observational or experimental research studies where data is captured for a subset of the population of interest, i.e. sample information. Well-designed research strives to ensure that the sample is representative of the population from which it is drawn. However, there is always a chance that the sample obtained is not representative of the population. Even if the sample is truly representative of the population, random error in the measurement can result in the captured data not being representative. If the analyst were to assume the data were perfectly representative, they would inappropriately present the certainty of the evidence base to the decision maker and hence increase the risk that a decision maker will inadvertently make the wrong decision.

### 4.2.2 *Extrapolation*

Many health care interventions have both positive and negative long-term effects. Often the length of time it takes for such effects to become clear is such that it is not feasible for them to be measured directly prior to the intervention being introduced into clinical practice. In the absence of direct evidence, the analyst is required to extrapolate from the observed data in order to capture an unbiased estimate of the true value of the intervention.

For example, consider a new treatment for breast cancer. The clinical trial reports recurrence and mortality rates for the untreated and treated patients with 2-year follow-up. At 2 years, there is a clear difference in recurrence favouring the new treatment (Fig. 4.1). Unless the decision maker's specifications require that outcomes are only considered over a shorter period, it would seem arbitrary and irrational to assume that the mortality and quality of life benefit of the new therapy will stop at the end of the trial follow-up period. Treating mortality and quality of life in this way has a similar effect to assuming that the 'additional' survivors on treatment die at the end of trial follow-up, even though there is no rational relationship between trial follow-up and survival or disease recurrence. To capture the full value of the new intervention compared to current practice, it will be necessary to extrapolate



**Fig. 4.1** Kaplan-Meier curve: disease relapse

from the advantage observed in a trial to predict how far that advantage will be maintained in the future. In Chap. 6 we consider a number of survival models that could be used for such extrapolation. The choice of model will impact upon the results of the analysis. Whilst it is not always necessary to produce results for alternative extrapolations, consideration of this source of uncertainty is important.

### 4.2.3 *Generalisability*

Generalisability is concerned with the degree of confidence with which data obtained in one setting can be used to inform a model of what should be expected in a different setting. Generalisability is a particularly acute concern when evaluating technologies close to launch, when almost all evidence is derived from the clinical trial development programme. Research studies, especially clinical trials for new therapies, tend to recruit atypical patients. Investigators want to be confident that any difference they observe between the experimental and control intervention can be attributed to the interventions rather than some difference in the characteristics of the patients in the two groups. For this reason complex patients tend to be excluded from clinical trials. The select nature of trial participants is further exacerbated by the select nature of the settings in which clinical trials are undertaken. Tertiary centres are much more likely to recruit patients into trials, and the care of patients in trials tends to be highly protocolised to further reduce intergroup variation. When a new technology is introduced into standard clinical care, the patients who receive the treatment and the care package within which the technology is delivered are likely to be substantially different to what had occurred in the trial. This creates some uncertainty about the degree to which the magnitude of differential effect seen in clinical trial will be seen in standard care. This is possibly the most obvious example of generalisability as a source of uncertainty. However, generalisability is an issue for other components of a cost effectiveness analysis (CEA) including quality of life/utility data, resource use and even the clinical practice inherent in the model structure.

### 4.2.4 *Model Structure*

A model structure attempts to describe the relevant clinical pathways for patients treated with the interventions being evaluated. As discussed in Sect. 3.2, the structure is by necessity a simplification of reality. The question is whether the simplifications exclude important characteristics of the potential care process that, if included in the model, would be likely to lead to a different decision? Whilst all models are wrong, some are useful (Box and Draper 1987), and the uncertainty associated with model structure is concerned with the risk that an excluded clinical pathway is associated with differences in costs and outcomes that could change the decision.

For example, consider an economic evaluation of non-steroidal anti-inflammatory drugs (NSAIDs) versus opioids in perioperative pain management. NSAIDs are associated with relatively small risks of major gastric bleeds. There is also a risk of death from major gastric bleeds. Managing gastric bleeds is very expensive, and a single death is likely to outweigh any population health benefit associated with avoiding opioid use. Should a cost effectiveness model examining the value of NSAIDs versus opioids for perioperative pain management include a gastric bleed health state, and should it include a gastric bleed-specific mortality risk?

The exposure to NSAIDs in the perioperative period is probably too short to induce a major gastric bleed. However, the relationship between duration of exposure and risk of bleed is an empirical question, and hence the justification for including or excluding gastric bleed and associated mortality from the model structure should be based upon the evidence on this issue. Failure to consider the issue explicitly and present the decision maker with the evidence for whatever choice is made will avoidably increase the decision uncertainty.

#### ***4.2.5 Methodological Uncertainty***

Methodological uncertainty exists in at least two levels in cost effectiveness models. First there is methodological uncertainty in the process of producing the evidence to parameterise the model; second there is uncertainty in the choice of modelling methods. As highlighted in Chap. 2, cost effectiveness models synthesise information from many different sources to produce predictions of future health outcomes and health care resource utilisation. Every component of these models is arrived at as the result of prior analyses underpinned by methodological choices, including:

- How should the effectiveness data from the trial be analysed?
- How should survival curves be extrapolated beyond the trial period?
- How should utility values be estimated for each health state included in the model?
- What model form should be used to predict future resource use?

These are all methodological choices, and different methods will often lead to different predictions and very different estimates of the uncertainty in those predictions. Often it is relatively straightforward to identify some methods as being absolutely incorrect; more rarely, it is sometimes possible to identify that only one reasonable approach remains. However, sometimes the choices are not so clear cut; which method to use is a matter of judgement, ideally supported by empirical tests of the face validity of the predictions obtained given directly observed data.

Methodological uncertainty in the choice of model method is primarily concerned with whether a decision tree, Markov, semi-Markov or patient-level simulation model is appropriate (choice of model is discussed in more detail in Chaps. 3 and 5). There is also increasing interest in the use of discrete event

simulation models for CEA.<sup>1</sup> In the context of decision tree, Markov and semi-Markov models, the choice of deterministic versus probabilistic is very important. So what do we mean by a deterministic model? A model is deterministic if its inputs are treated as certain. Remember the decision tree we constructed in Chap. 3? This was a deterministic model; we did not incorporate uncertainty directly into the analysis. But, we can use the model to explore uncertainty in terms of how sensitive outputs are to changes in certain inputs. The model is deterministic as it does not directly describe the uncertainty in the inputs or outputs. The challenge with deterministic models is that they will produce biased estimates of non-linear processes. A probabilistic model incorporates uncertainty within the model and is typically referred to as a stochastic model. Within Markov models, the parameters are stochastic (i.e. there is uncertainty about their value within an analysis). Probabilistic models are discussed in Sect. 4.3 of this Chapter.

In respect of choice of model beyond the factors identified here, there are few clear rules for identifying the appropriate modelling method, but different model forms can produce different results (Brennan et al. 2006). It is important that the analyst be explicit about the rationale for the modelling strategy adopted, ensuring that the decision maker is aware of any uncertainty that this introduces.

### 4.3 Analytic Responses to Uncertainty in CEA

Sampling variation is typically addressed through the calculation of 95 % confidence intervals. In the classical statistics framework, the confidence intervals around a mean value estimate the range within which a sample mean would lie for 95 samples out of 100 if they were sampled from the same population. Confidence intervals are often interpreted, incorrectly, as a 95 % subjective probability that the population mean lies in the interval range. It is worth noting that subjective probabilities are a Bayesian concept and the Bayesian equivalent of the confidence interval is instead called a credible range.

Confidence intervals are problematic for cost effectiveness analyses, because the ICERs are not statistically well behaved because they are formed as a ratio of two uncertain variables. Further, as there are multiple ways of achieving the same ratio value, the ICER of an intervention versus its next best alternative does not normally have the distributional characteristics assumed by parametric methods for calculating confidence intervals. A number of authors have proposed approximations for classical confidence intervals for ICERs, although none of these are wholly satisfactory (Briggs and Fenn 1998).

---

<sup>1</sup> Discrete event simulation modelling is beyond the scope of this text. Karnon (2003) provides a comparison of Markov and discrete event simulation modelling in cost effectiveness. More recently Caro et al. (2010) argue for the use of discrete event simulation in preference to Markov modelling.

**Table 4.1** Type of sensitivity analysis

Type of sensitivity analysis	Definition
One way	The variation of one parameter whilst all other parameters are held constant to observe the impact on the predicted costs and outcomes
Multiway	The variation of more than one parameter whilst all other parameters are held constant to observe the impact on predicted costs and outcomes
Threshold analysis	The variation of a one parameter to identify the value at which the decision implications of the costs and outcomes change
Analysis of extremes	Setting one or more parameters at their extreme (upper and lower) values to observe the impact on predicted costs and outcomes
Probabilistic sensitivity analysis	Using probability distributions to characterise the credible range and the likelihood of any given value being observed, combined with simulation to promulgate the uncertainty in the parameters to uncertainty in the predicted costs and outcomes

In the absence of a direct equivalent to the 95 % confidence interval that many decision makers are comfortable with, analysts have developed a number of methods for exploring the uncertainty in cost effectiveness model results. These methods fall under the umbrella term *sensitivity analysis*. Sensitivity analysis is the process of varying model input values and recording the impact of those changes on the model outputs. Five different types of sensitivity analysis are reported in the literature: one-way, multiway, threshold, analysis of extremes and probabilistic sensitivity analysis. These are summarised in Table 4.1, and each is described in more detail above.

### 4.3.1 One-Way Sensitivity Analysis

*One-way sensitivity analysis* examines how the ICER changes in response to changes in a single parameter whilst holding the value of all other model parameters constant. It is frequently used to establish that a model is constructed correctly by testing whether the direction of change in the model results is that logically required by the direction of change in the parameter that is varied. It has also been used to identify which parameters in a model have the strongest direct effect on the model results and hence have particular importance for the decision maker. This use of one-way sensitivity analysis is discussed in more detail below when we consider probabilistic sensitivity analysis. The great limitation of one-way sensitivity analysis is the implication of keeping all other parameters fixed. This assumes that there is no causal relationship between the value taken by one parameter and the value taken by other parameters. Only a little reflection is required to conclude that this is unlikely to be the case. For example, if a treatment that aims to improve quality of life is more effective than expected, then this may be because there have been fewer side effects than expected, which require less additional treatment, fewer resources and so reduced costs. Similarly, if medical staffing costs are higher, this may reflect more experienced and hence more effective medical staff inputs, which should be

**Table 4.2** Two-variable example of multiway sensitivity analysis

	X -95CI (lower bound of 95 % CI)	X mean	X +95CI (upper bound of 95 % CI)
Y -95CI (lower bound of 95 % CI)	1	2	3
Y mean	4	5	6
Y +95CI (upper bound of 95 % CI)	7	8	9

reflected in the model effectiveness parameter. The arbitrary variation of the value of one particular parameter provides information that is easy to misinterpret. As a result, one-way sensitivity analysis is increasingly recognised as a limited tool for the analysis of uncertainty. However, it is useful as a mechanism for debugging models.

### 4.3.2 *Multiway Sensitivity Analysis*

Multiway sensitivity analysis varies more than one parameter at once and examines the impact of the different combinations of changes on the model outputs. A regularly used strategy is to vary parameters, allowing each parameter to take both its mean value and those of its upper and lower confidence limits. Table 4.2 shows how, considering only two parameters ( $X$  and  $Y$ ), using this approach to multiway sensitivity analysis produces nine different sets of results.

Whilst it is feasible for a decision maker to examine nine different results, the multiway sensitivity analysis rapidly become infeasible as the number of parameters in a model increases. For example, a model with ten parameters would create over 59,000 different sets of results if each parameter is varied using the mean, upper and lower 95 % confidence interval value.

A further problem with multiway sensitivity analysis is that it treats all possible combinations of the values as equally valid. The underlying relationship between the parameters described above means that this is unlikely to be true. Some combinations of parameter values are likely to be in conflict with the underlying relationships. As a result, multiway sensitivity analysis risks misleading decision makers as much as informing them, unless the specific combinations of parameter values are assessed for their face validity and relevance to the decision problem. When the set of parameter values is chosen for a specific reason, this is called scenario analysis. Decision makers are often interested in worst- and best-case scenarios or specific policy scenarios that are not necessarily founded in the evidence base. Hence, multiway sensitivity analyses continue to be viewed as useful mechanisms for exploring uncertainty.

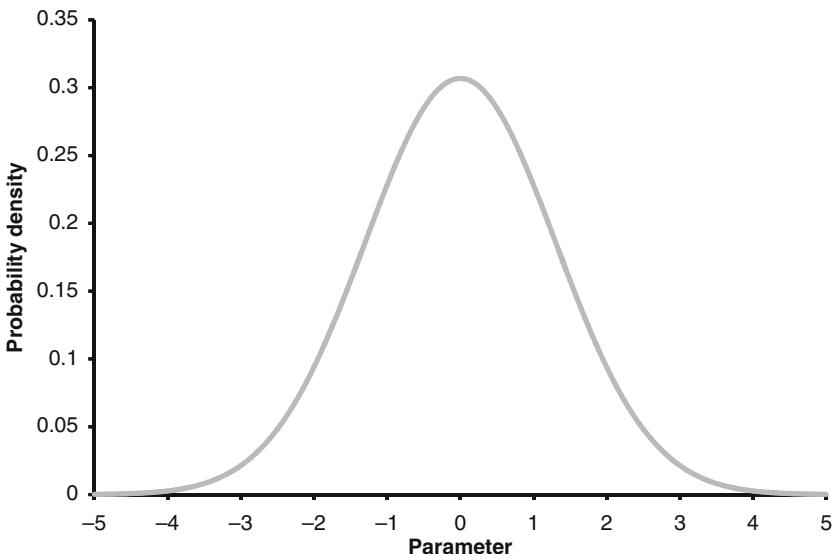
### 4.3.3 Threshold Analysis

Threshold analysis considers what value a specific parameter must take in order to achieve a specified change in the model results or to achieve a target result; for example, a single parameter might be manipulated to set an ICER equal to the ceiling ratio. Decision makers may also request threshold analyses around effectiveness or safety parameters to inform their deliberations on the likelihood that a technology could be cost effective. However, threshold analyses are a specific form of one-way sensitivity analysis and thus ignore any underlying relationships between the parameter being varied and the value taken by other parameters in the model. Hence analysts should provide simplistic threshold analyses results to decision makers with very strong caution about the inherent risk of error from such analyses.

### 4.3.4 Analysis of Extremes

Analysis of extremes considers the impact of setting one or more parameters at the highest or lowest possible value. Depending upon whether this is done for one or multiple parameters, the problems described above make such analyses fraught with difficulties for interpretation. A possible exception to this is where the price of the intervention is set at its extreme value. When the characteristics of the intervention are not affected by its price – e.g. a new drug – then it is valid to use analysis of extremes to consider whether the technology would be cost effective at the lowest possible price.

All these forms of sensitivity analysis assume that the parameters take a single known value and change this value as a means of exploring the uncertainty in the model results. As described in Sect. 4.2.5, models where each parameter has a single known value are referred to as *deterministic cost effectiveness models*. The alternative is to recognise that we are uncertain about the true value of almost all parameters in a model and incorporate this uncertainty into the model directly. This is done by replacing the single values of the deterministic model with probability distributions, reflecting the expected value and the uncertainty around that expectation for each parameter. These models are referred to as *stochastic cost effectiveness models*. The exploration of uncertainty in a stochastic cost effectiveness model is done through *probabilistic sensitivity analysis* (PSA). PSA combines the probability distributions for each of the input parameters with *Monte Carlo simulation* to produce probability distributions for each of the outputs of the model, hence providing a quantification of the credible range for the expected ICER.



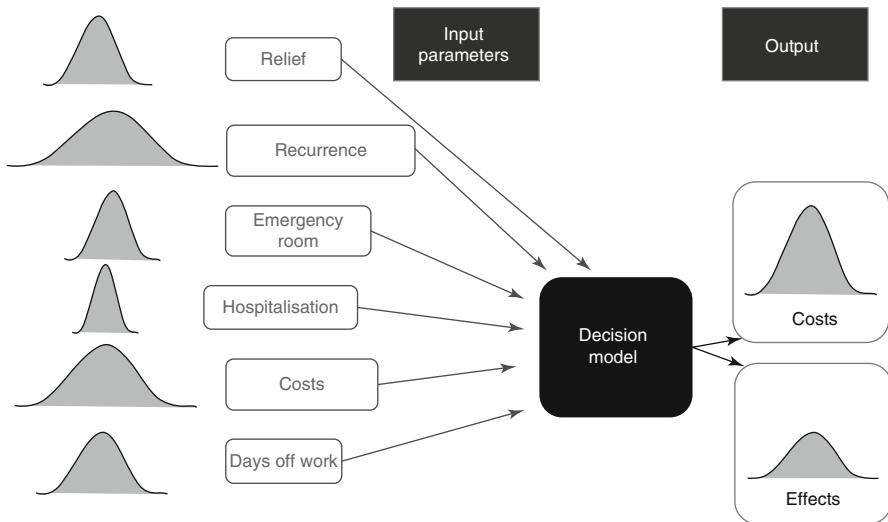
**Fig. 4.2** Probability distribution

#### 4.4 Probabilistic Sensitivity Analysis (PSA)

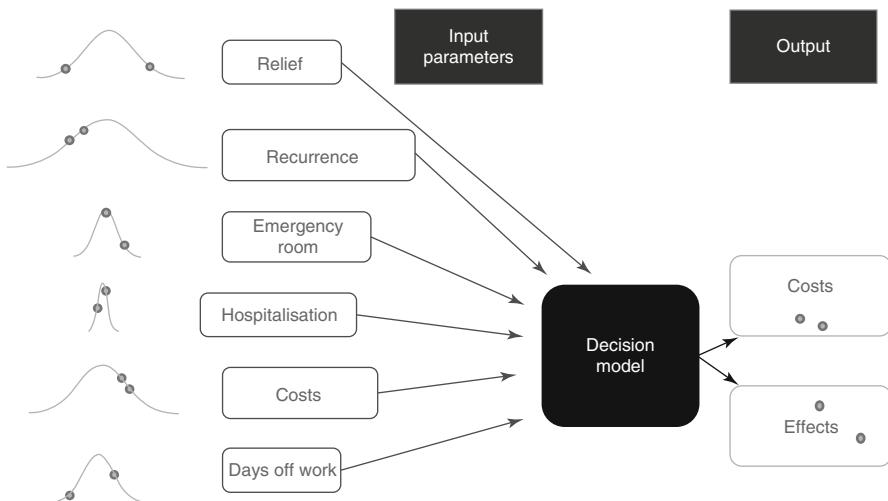
PSA requires that each parameter is expressed as a probability distribution. Each input to the model is known with a certain degree of uncertainty. This uncertainty can be characterised as a probability density function with an associated probability distribution. Figure 4.2 shows a probability distribution for a parameter that has an expected value of zero and an effective range of plus four to minus four. The parameter is normally distributed, and hence strictly the range is plus/minus infinity, and the probability is symmetrical, meaning that values equidistant from the expected value are equally likely.

Once a probability distribution has been described for each parameter in the cost effectiveness model, PSA allows the analyst to vary the value of each parameter simultaneously. Values are drawn from the probability distributions, and the outputs of the model for each draw are recorded. The specification of the probability distribution is based upon observed data for the parameters included in the model. By repeating the process of draws from the parameter probability distributions and capturing the outcome predictions, probability distributions for the costs and outcomes for the technologies being compared are obtained.

It is important to note that the distributions for the costs and outcomes of each technology are the primary outputs from probabilistic sensitivity analysis. The ICER is then constructed using the mean values of each of these distributions. Whilst each draw allows the construction of a draw specific health ICER, if the



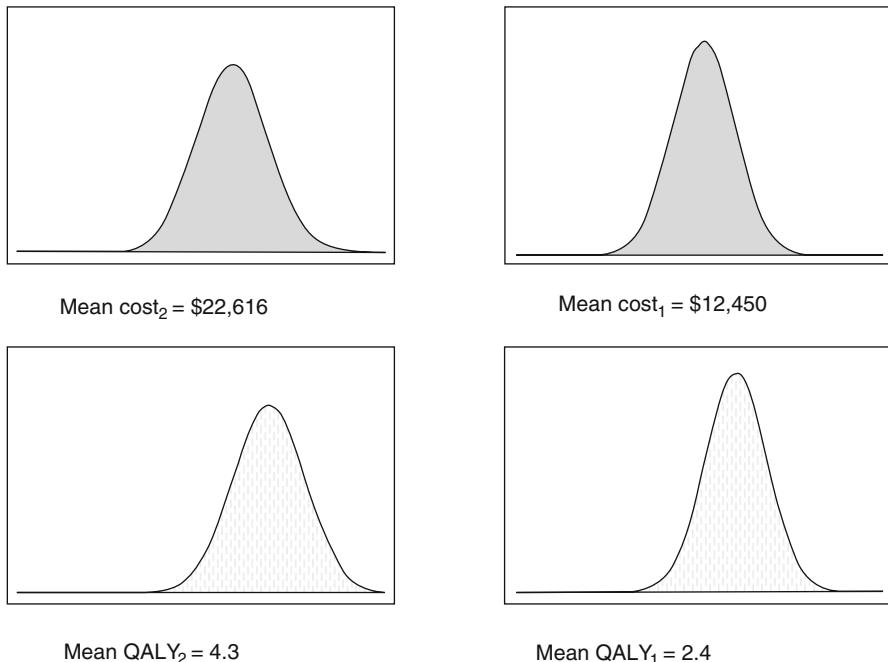
**Fig. 4.3** Parameter inputs for a cost effectiveness model for migraine



**Fig. 4.4** Parameter inputs for a cost effectiveness model for migraine with two random draws from each distribution

distribution of such estimates was to be calculated, the mean of that distribution would not be the appropriate estimate of the cost effectiveness of the technologies. This is because the mean of a ratio is not equal to the ratio of the means. The unbiased estimate of the ICER is calculated using the means of the individual distributions, as illustrated in Fig. 4.3. Figure 4.4 illustrates the concept of sampling to produce the results of a stochastic cost effectiveness model.

Model parameters include probability distributions for transition to the other states, costs and quality of life weights. A selection of these parameter distributions is illustrated on the left-hand side of the figure. We have highlighted two random draws from each of these distributions, which are used in the decision model to produce two separate predictions of costs and effects, shown on the right-hand side of the figure. When a sufficient number of draws have been made, usually several thousands, it is possible to construct probability distributions for the costs and outcomes as illustrated in Fig. 4.5, from which we use the expected values to calculate the ICER (Box 4.1).



**Fig. 4.5** Probability distributions for costs and outcomes

#### Box 4.1: ICER

$$\text{Expected incremental cost} = C_2 - C_1 = \$22,616 - \$12,450 = \$10,166$$

$$\text{Expected incremental effects} = E_2 - E_1 = 4.3 - 2.4 = 1.9 \text{ QALYs}$$

$$\text{ICER} = \$10,166 / 1.99 = \$5,351 \text{ per QALY gained}$$

## 4.5 Outputs from Probabilistic Analysis

It is important to understand that a probabilistic analysis will produce an estimate of the ICER for each individual run of the model. However, the best estimate of the ICER is not provided by the mean of the distribution of these ICERs. The best prediction of the ICER is constructed by using the expected value of the cost for each intervention and the expected value of the outcome for each intervention. This point is often summarised as the ‘mean of the ICER’ does not equal the ‘ICER of the means’, and it is the ICER of the means we need to report to decision makers. Exercise 4.1 is designed to help you to understand the difference between the two measures, as it provides the data from a small sample of simulations to calculate both measures and see how they differ.

### Exercise 4.1 Exploring a Stochastic Cost Effectiveness Model

The aim of this exercise is to examine the difference between the expected incremental cost effectiveness ratio calculated as the ratio of the incremental means compared to the mean of the ratios. In Table 4.3 we have reported the outputs, costs and QALYs, for a sample of ten sets of draws, for two alternative treatments.

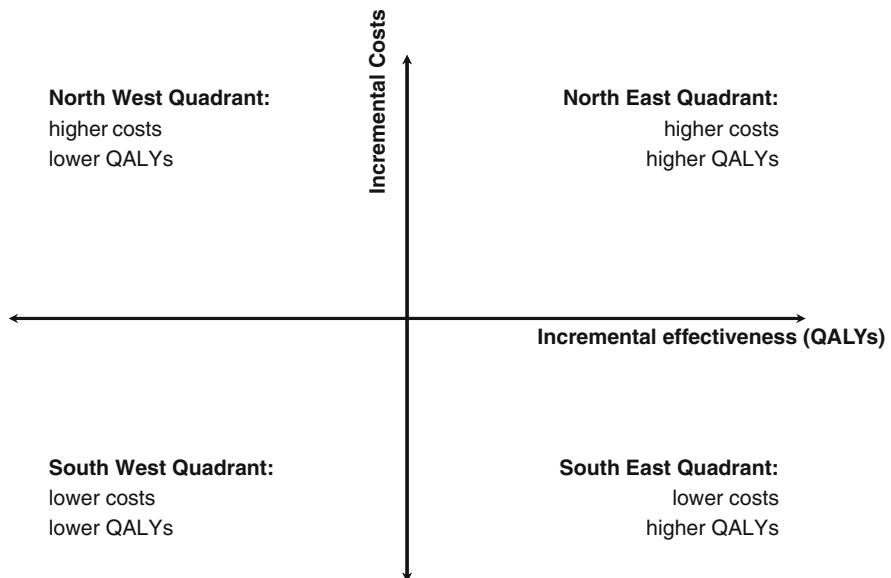
First calculate the incremental cost effectiveness ratio for each draw, and then calculate the mean of these ratios.

Second calculate the mean costs for Intervention 1 and Intervention 2 and then the mean QALYs for Intervention 1 and Intervention 2. Use these means to estimate the expected ICER.

1. Are the two results the same?
2. Can you think of why the second method of calculation is the correct estimation of the ICER?
3. MS Excel (R) worksheets for this exercise can be downloaded from <http://medhealth.leeds.ac.uk/costeffectivenessmodelling>

**Table 4.3** Costs and QALYs for a sample of ten draws

Simulation run	Cost_1	QALYs_1	Costs_2	QALYs_2	ICER
1	\$111	0.0019	\$130	0.0023	
2	\$66	0.0021	\$64	0.0026	
3	\$45	0.0020	\$35	0.0022	
4	\$75	0.0015	\$81	0.0021	
5	\$70	0.0017	\$74	0.0024	
6	\$43	0.0013	\$35	0.0019	
7	\$93	0.0018	\$109	0.0024	
8	\$52	0.0017	\$44	0.0023	
9	\$75	0.0017	\$85	0.0025	
10	\$55	0.0014	\$54	0.0021	
Mean					Expected ICER



**Fig. 4.6** The cost effectiveness plane

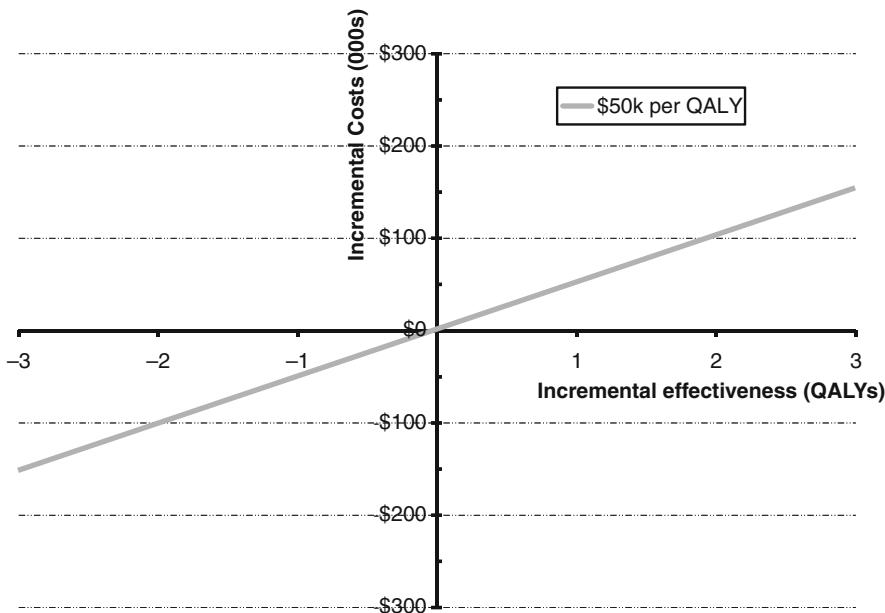
## 4.6 Some Problems with ICERs

It is clear that ICERs provide a neat and easily accessible way in which to demonstrate the decision rules of cost effectiveness analysis (Karlsson and Johannesson 1996). However, they do have limitations; we can think through some of these limitations using the cost effectiveness plane. As described in Sect. 1.4, the cost effectiveness plane is the two-dimensional space in which we can plot all possible values for the ICER. Incremental costs are measured on the vertical axis, and incremental outcomes, normally QALYs, are measured on the horizontal axis.

Figure 4.6 shows the cost effectiveness plane. The plane consists of four quadrants, and these are conventionally discussed using the compass terminology of the North West (top right-hand quadrant), South West (bottom right-hand quadrant), South East (bottom left-hand quadrant) and North East (top left-hand quadrant).

In the North West quadrant are the ICERs for all technologies that have positive incremental costs and positive incremental QALYs. Adopting these technologies will either require an increase in the health care budget or displacement of other technologies that are currently funded from the budget.

In the South East quadrant are the ICERs for all technologies that have negative incremental costs combined with positive incremental QALYs. These are the technologies every budget holder would love to see as they increase total health and whilst freeing up resources. As described in Sect. 1.4.1, such technologies are described as ‘dominant’. In the South West quadrant, we observe the ICERs for technologies that are cheaper and less effective. To understand whether these technologies could represent good value requires insight into whether the savings could



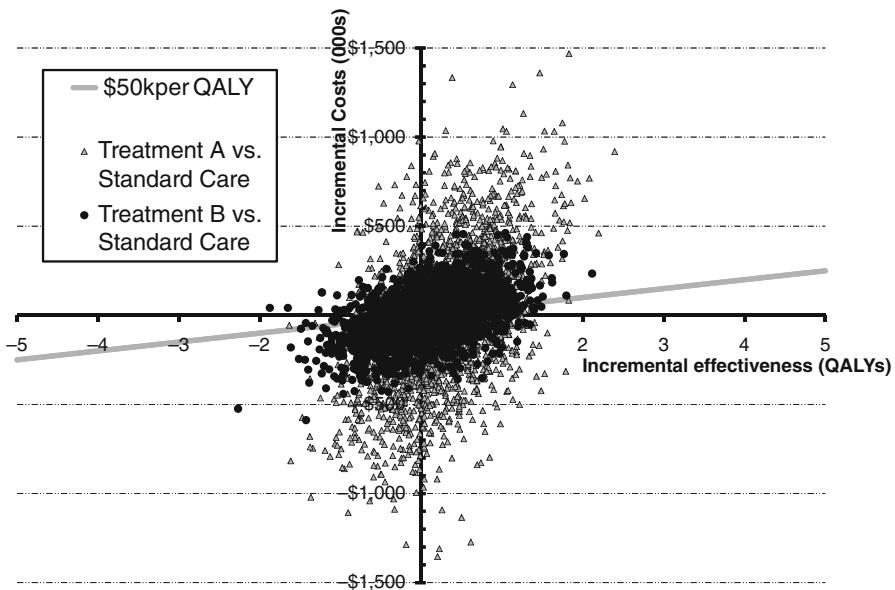
**Fig. 4.7** A cost effectiveness threshold on the cost effectiveness plane

produce more health through reallocation to other activities than is lost through their implementation. We will discuss this in more detail below.

Finally in the North West quadrant, we observe the ICERs for those technologies that both cost more money and produce less health. These technologies are said to be ‘dominated’. The adoption of dominated technologies will unequivocally decrease the health produced by the health system.

The cost effectiveness threshold can be represented on the cost effectiveness plane by a straight line passing through the origin with the appropriate slope. Hence, a \$50,000 per QALY threshold is plotted by a line that passes through the points (50,000, 1) and (-50,000, -1) (Fig. 4.7). In the context of a fixed health system budget, technologies with ICERs to the left of the threshold line are ‘cost effective’; i.e. their adoption will be expected to increase population health, whilst technologies to the right of the threshold are not cost effective, and their adoption will be expected to reduce population health (Fig. 4.7).

We can use the cost effectiveness plane to represent the uncertainty in the cost effectiveness results by plotting the ICERs for each simulation on the plane. A scatter plot on the cost effectiveness plane for a technology with a high level of certainty regarding the actual ICER will be tightly clustered data points, whilst increasing uncertainty will be shown by increasingly dispersed data points. Decision uncertainty is shown in the degree to which the data points are to the left or the right of the cost effectiveness threshold. Whilst the scatter plot gives a qualitative sense of the scale of uncertainty, with literally thousands of simulations plotted on the plane, confident interpretation can be difficult. For example, consider an economic evaluation that compares three technologies. It is possible, even likely, that the scatter



**Fig. 4.8** Scatter plot on the cost effectiveness plane: multiple technologies

plots will overlap, as shown in Fig. 4.8. We have two new technologies ( $NT_1$  and  $NT_2$ ) being compared to standard care (SC). The triangular markers are simulated ICERs for  $NT_1$  vs. SC, and the circular markers are simulated ICERs for  $NT_2$  vs. SC. By examining Fig. 4.8, could you comment as to whether we are more uncertain as to the cost effectiveness of  $NT_1$  or  $NT_2$  compared to SC?

In plotting the cost effectiveness threshold on the plane, we can see that ICERs are not unique; every point on the threshold line has the same value. In fact there are an infinite number of combinations of costs and effects that will produce the same ICER. Most importantly, you can get the same ICER for a technology that has positive incremental costs and incremental QALYs and one that produces negative incremental costs and QALYs. Equal ICERs do not mean equal value. ICERS in the North East and South West quadrants must be interpreted differently. We might consider this a structural uncertainty in the interpretation of ICERs. Fortunately, we can use the cost effectiveness threshold to overcome this problem by converting an ICER into an alternative statistic, the Net Benefit (Stinnett and Mullahy 1998).

As described in (Sect. 1.4.2), Net Benefit can be expressed in monetary or health terms. Net Monetary Benefit (NMB) is calculated by dividing the QALYs by the value of the cost effectiveness threshold ( $\lambda$ ) and then subtracting the cost from the result. It is the monetary value of the additional benefit of the new technology after the incremental costs have been netted out. Net Health Benefit (NHB) is calculated by dividing the cost of the new technology by the threshold ( $\lambda$ ) and subtracting the result from the QALYs produced. The intervention with the higher Expected Net Benefit is the cost effective choice. The differences in the Expected Net Benefit measures the expected health gain from the new technology once the health displaced through opportunity cost has been netted out, per person treated.

**Table 4.4** Net Monetary and Net Health Benefit calculations

Cost	Incremental QALYs	Incremental costs/ $\lambda$	Net Health Benefit	QALYs * lambda	Monetary Benefit	$\lambda$
\$23,432	4.74	0.47	4.27	\$213,568	\$190,136	\$50,000
\$18,318	3.02	0.37	2.65	\$132,682	\$114,364	
\$21,010	3.36	0.42	2.94	\$146,990	\$125,980	
\$16,840	2.51	0.34	2.17	\$108,660	\$91,820	
\$19,333	2.77	=A6/\$G\$2	=B6-C6	=D6*\$G\$2	=E6-A6	

Table 4.4 has some exemplar incremental costs and QALYs and the associated Net Health and Net Monetary Benefit, assuming that  $\lambda$  is valued at \$50,000. In the final row, we have included the Excel formulas for the calculation. You might want to check whether you get the same results, to confirm your understanding. In the formula cell G2 provides the value for lambda, costs and incremental QALYs appear in columns A and B.

Note that the uncertainty we observe in the scatter plot on the cost effectiveness plane is maintained in the Net Benefit framework, whilst the problem of nonunique ICERs is eliminated. We can calculate the NHB and NMB for each simulation. In the example shown in Figure 4.8, we could calculate the proportion of simulations where the NHB for SC, NT<sub>1</sub> and NT<sub>2</sub> is positive and the proportion of simulations when:

- A. NHB NT<sub>1</sub>>SC
- B. NHB NT<sub>2</sub>>SC
- C. NHB NT<sub>2</sub>>NT<sub>1</sub>

## 4.7 Summary

- Uncertainty is a key concern for analysts in decision analytic cost effectiveness modelling. The analyst needs to consider whether the nature or the magnitude of the uncertainty is such that there is a substantial risk making the wrong decision using the available information.
- There are five potential sources of uncertainty in modelling: sampling variation, extrapolation, generalisability, model structure and methodological uncertainty.
- Sampling variation refers uncertainty over whether the sample is representative of the population.
- The analyst often has to extrapolate from the observed data in order to capture long-term effects of many health care interventions; this extrapolation is often a source of uncertainty.
- Generalisability is concerned with the degree of confidence (certainty) with which data obtained in one setting can be used to inform a model of what should be expected in a different setting.
- The uncertainty associated with model structure is concerned with the risk that an excluded clinical pathway is associated with differences in costs and outcomes that could change the decision.

- There is methodological uncertainty in the process of producing the evidence to parameterise the model and in the choice of modelling methods.
- Sensitivity analysis is the process of varying model input values and recording the impact of changes on the model outputs. Types of sensitivity analysis include one-way, multiway, threshold, analysis of extremes and probabilistic sensitivity analysis.
- One-way sensitivity analysis is the variation of one parameter, whilst all other parameters are held constant to observe the impact on the predicted costs and outcomes.
- Multiway sensitivity analysis involves the variation of more than one parameter, whilst all other parameters are held constant to observe the impact on predicted costs and outcomes.
- Threshold analysis involves the variation of one parameter to identify the value at which the decision implications of the costs and outcomes change.
- Analysis of extremes sets one or more parameters at their extreme (upper and lower) values to observe the impact on predicted costs and outcomes.
- Probabilistic sensitivity analysis uses probability distributions to characterise the credible range and the likelihood of any given value being observed, combined with simulation to propagate the uncertainty in the parameters to uncertainty in the predicted costs and outcomes.
- The cost effectiveness plane may be used to represent uncertainty in the cost effectiveness results by plotting the ICERs for each simulation on the plane. However, there are an infinite number of combinations of costs and effects that will produce the same ICER. We can use the cost effectiveness threshold to convert an ICER into NMB or NHB.
- The technology with the highest NMB or NHB is the highest value technology.

## References

- Box GEP, Draper NR (1987) Empirical model building and response surfaces. Wiley, New York
- Brennan A, S C, Davies R (2006) A taxonomy of model structures for economic evaluation of health technologies. *Health Econ* 15(12):1295–1310
- Briggs AH, Fenn P (1998) Confidence intervals or surfaces? Uncertainty on the cost-effectiveness plane. *Health Econ* 7(8):723–740
- Caro JJ, Möller J, Getsios D (2010) Discrete event simulation: the preferred technique for health economic evaluations? *Value Health* 13(8):1056–1060. doi:[10.1111/j.1524-4733.2010.00775.x](https://doi.org/10.1111/j.1524-4733.2010.00775.x)
- Karlsson G, Johannesson M (1996) The decision rules of cost effectiveness analysis. *Pharmacoeconomics* 9(2):113–120
- Karnon J (2003) Alternative decision modelling techniques for the evaluation of health care technologies: Markov processes versus discrete event simulation. *Health Econ* 10(12):837–848
- Stinnett A, Mullahy J (1998) Net health benefits: a new framework for the analysis of uncertainty in cost effectiveness analysis. *Med Decis Making* 18(2-suppl):S68–S80

# Chapter 5

## Introduction to Markov Cost Effectiveness Models

**Abstract** Modelling in cost effectiveness provides a common framework by which to systematically assess costs and outcomes, to represent uncertainty and to compare alternative courses of action. Having introduced decision trees for cost effectiveness modelling, we now move on to Markov models. This chapter provides an introduction to Markov models for cost effectiveness analysis, guiding you through when Markov modelling should be used; the concept of health states (which for the backbone or structure of a Markov model) and transition probabilities, Markov trace and cycles, time horizon and discounting.

### 5.1 Introduction

Recall in Chap. 3 we said that modelling in cost effectiveness provides a common framework by which to systematically assess costs and outcomes, to represent uncertainty and to compare alternative courses of action and that, more formally, *decision modelling can be defined as a systematic approach to decision making under conditions of uncertainty, in which the probability of each possible event, along with the consequences of those events, is explicitly stated* (Kielhorn and von der Schulenburg 2000). Having introduced decision trees for cost effectiveness modelling in Chap. 3, we now move on to Markov models. Markov models are typically more complex than decision trees. They enable us to address a wide range of questions, and, unlike most decision trees, they can explicitly take account of time.

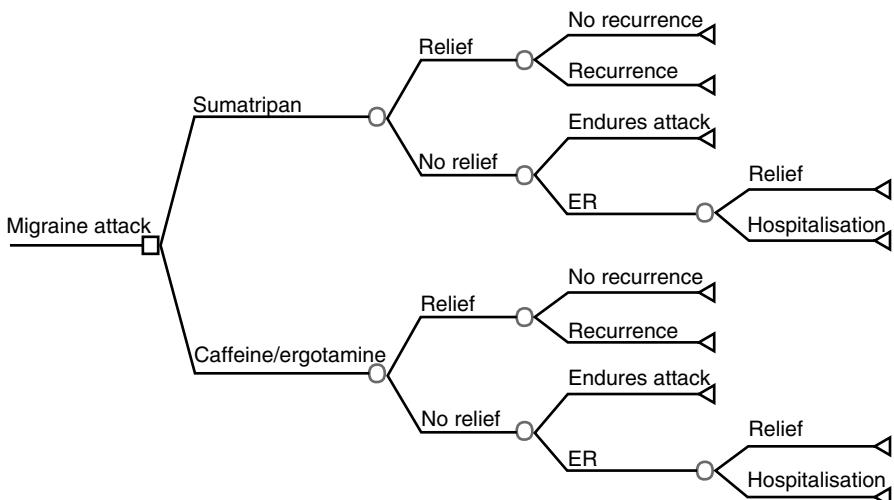
Chapter 5 provides an introduction to Markov models for cost effectiveness analysis and is structured as follows: Section 5.2 discusses when Markov modelling should be used. Section 5.3 describes the concept of health states which form the backbone or structure of a Markov model. The concepts of transition probabilities, Markov trace and cycles, time horizon and discounting complete the Chapter in Sects. 5.4, 5.5 and 5.6, respectively, followed by a summary of key points in Sect. 5.7.

## 5.2 Why Use Markov Models?

As highlighted in Chap. 3, the decision to use modelling within cost effectiveness analysis is driven by a number of factors including the limitations and availability of single trial data to address the question you are asking (see Box 3.1). These factors are equally relevant to the use of Markov modelling in cost effectiveness analysis as they are to decision trees. Similarly the components of a good economic model also apply: models should be populated with the most appropriate and good quality clinical data, reflect a realistic picture of current clinical practice, use the appropriate comparator(s) be run for an appropriate time period, be valid, transparent and reproducible, explore uncertainty and be easily interpreted.

Once we have decided that a model is appropriate for our cost effectiveness analysis, we need to decide on the type of model. Markov models are most useful when health events repeat over time or have longer-term health effects; when the effect of treatment either stops quickly after an initial treatment or continues at its earlier level; and when the risk of different health events does not depend on the patient's prior history. Decision trees are limited in as much as they are designed to capture what happens *at a point in time*; there is no explicit sense of time passing. Unlike most decision trees, Markov models enable us to incorporate the passage of time.

However, decision trees are good for considering transitory health conditions. Imagine we want to model a condition which has regular flare-ups that can require hospitalisation such as migraines. This might be modelled using a decision tree like that shown in Fig. 5.1 (you might recognise this from the exercise in Chap. 3!).



**Fig. 5.1** Decision tree: sumatriptan vs. caffeine/ergotamine for migraine attacks (Evans et al. 1997)

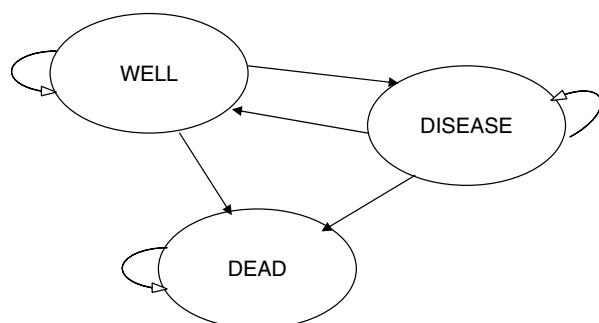
But what alternatives are open to us when decision models became very complex – when the trees became *bushy*? Even if we've only got three possible outcomes giving three nodes at the end of the first flare-up and nine nodes at the end of the second flare-up; using a decision tree, by the end of the tenth cycle, we would have 59,049 nodes! The way that Markov models are structured gets round this problem as the model has health states that individuals can transition between (forwards and backwards); whereas decision trees are unidirectional, the individual can only move from left to right in the model.

It is important to note at this point that whilst Markov models and decision trees differ, the thought process and planning in terms of things we need to think about *prior* to building the model are the same. We need to formulate the decision problem, the question we want to ask. The question should involve at least two options or interventions, costs and at least one outcome upon which to base a recommendation. The perspective the evaluation takes should be determined and the target population identified. Once again it is imperative that we have a good understanding of the natural history of the health condition that we are interested in. Once all these things have been thought through, you can begin the task of building the model.

### 5.3 Health States

The natural history of the condition is very important in designing a Markov model given that the model is structured around health states and movements between them. Within Markov models, health is split into distinct categories. These categories or health states must be mutually exclusive and cover all the people in the model (everyone must fit into a health state at any point in time). Individuals can only be in one state at a time and will stay in that state for a specified or fixed period of time. This period of time is known as a cycle, and at the end of each cycle, the patient can stay in the same health state or move to another health state.

Imagine a very simple model in which we have three health states: Well, Disease and Dead. We can show the health states and movements between them in the form of a picture known as an influence diagram. Our simple model is shown below (Fig. 5.2).



**Fig. 5.2** Simple Markov model

We can see that the three health states are mutually exclusive. Patients can only be well, have the disease or be dead within the model. Individuals in the model move at the end of each period or cycle along an arrow. For example, if you start in the Well health state, you can move to the Disease state or the Dead state or you can stay in the Well state. Similarly you can move from the Disease state to the Well state or to the Dead state or remain in the Disease state. Once in the Dead state, you can't move from there to any other state – once you're dead, you stay dead!

Although influence diagrams are not generally used to do calculations, they are very useful. They are a great source of visual information to help us to understand the pathway and potential impact of the condition of interest. We can see how people transit between health states using the arrows connecting the states. However, once we understand the potential patient journeys, we need more information than the current diagram can provide. When thinking about the patient journey, we need to think about how long we need to follow a process for, how long people stay in each state, how often people stay in each state and what costs and outcomes are associated with being in each.

## 5.4 Transition Probabilities

Let's consider the movements between states that are represented by the arrows in the influence diagram. What is the chance of an individual who is well contracting the specified disease represented in the model, i.e. moving from the Well health state to the Disease health state? In answering this question we need to know the probability of moving between the two states – this is known as the transition probability. Transition probabilities are a key element of a Markov model; the transition probability predicts how people will move from one health state to another. The probability of moving to disease at the end of period, given that you started in the period  $t$  in Well, is summarised as  $P(\text{Disease}_{t+1} | \text{Well}_t)$ . Transition probabilities can be written on the tree (alongside the relevant arrows), but more often they are presented in the form of a transition matrix. An example is shown in Table 5.1.

The Transition Matrix is a way of combining conditional probabilities together. Thinking about the transition between health states, what can we say about the transition  $P(\text{Well}_{t+1} | \text{Dead}_t)$ ? Quite simply this will always have a transition probability of zero. If a person is Dead in time period 1, then he/she cannot transition to Disease in the next time period. Table 5.2 gives the transition in figures. You can see from the matrix that entries in each row sum to one.

The Markov *assumption* states that you use the same transition matrix every time; however, this assumption is relaxed for transition changes by time. A Markov

**Table 5.1** Transition probabilities for our simple Markov model (1)

		End of period $t$ (start of period $t+1$ )		
		Well	Disease	Dead
Start of period $t$	Well	$P(\text{Well}_{t+1}   \text{Well}_t)$	$P(\text{Disease}_{t+1}   \text{Well}_t)$	$P(\text{Dead}_{t+1}   \text{Well}_t)$
	Disease	$P(\text{Well}_{t+1}   \text{Disease}_t)$	$P(\text{Disease}_{t+1}   \text{Disease}_t)$	$P(\text{Dead}_{t+1}   \text{Disease}_t)$
	Dead	$P(\text{Well}_{t+1}   \text{Dead}_t)$	$P(\text{Disease}_{t+1}   \text{Dead}_t)$	$P(\text{Dead}_{t+1}   \text{Dead}_t)$

**Table 5.2** Transition probabilities for our simple Markov model (2)

		End of period $t$ (start of period $t+1$ )		
		Well	Disease	Dead
Start of period $t$	<b>Well</b>	0.7	0.2	0.1
	<b>Disease</b>	0.1	0.6	0.3
	<b>Dead</b>	0	0	1

model can allow the transition matrix to change; the transition probabilities can depend on the period, but the same figures apply for everyone.

## 5.5 Markov Trace

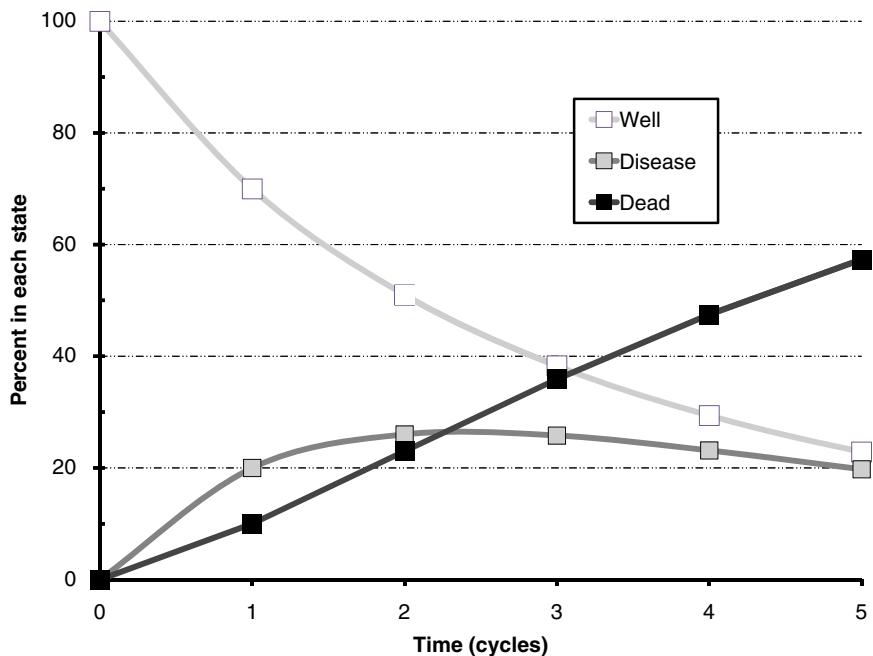
Having looked at transition probabilities to estimate the chance that people will move between health states or remain in their current health state – we can now go on to think about how long people stay in each state and how long we need to follow the process for.

The transition matrices are very mechanistic. Applying the transition matrix, each period or cycle is repeated over and over again, each time we move the people within the model (the cohort) on another period of time. The Markov trace captures the numbers or proportion of people in each health state in each time period and how that changes over time. The Markov trace uses the transition probabilities to calculate the movement between groups. Using our previous example (Table 5.2), imagine we begin in time period 1 with a cohort of 100 people in the Well health state. No one is in the Disease or Dead states. At the start of period  $t+1$ , we will have 70 people who have remained well, 20 people who have contracted the disease and 10 people who have died. At the start of the next period ( $t+2$ ), of the 70 in the Well health state, 49 will remain there, 14 will move to the Disease state and 7 to the Dead state. Of the 20 people in the Disease state, 2 will have moved to the Well state, 12 will remain in the Disease state and 6 to the Dead state. This means that at  $t+2$ , we now have 51 people in the Well state, 26 in the Disease state and 23 who have died. This is repeated over the time horizon of the model. Table 5.3 traces the transitions over the first 6 time periods. We can see that by time period 5 well over 50 % of the cohort are in the Dead state. Figure 5.3 shows our Markov trace for the figures in Table 5.3.

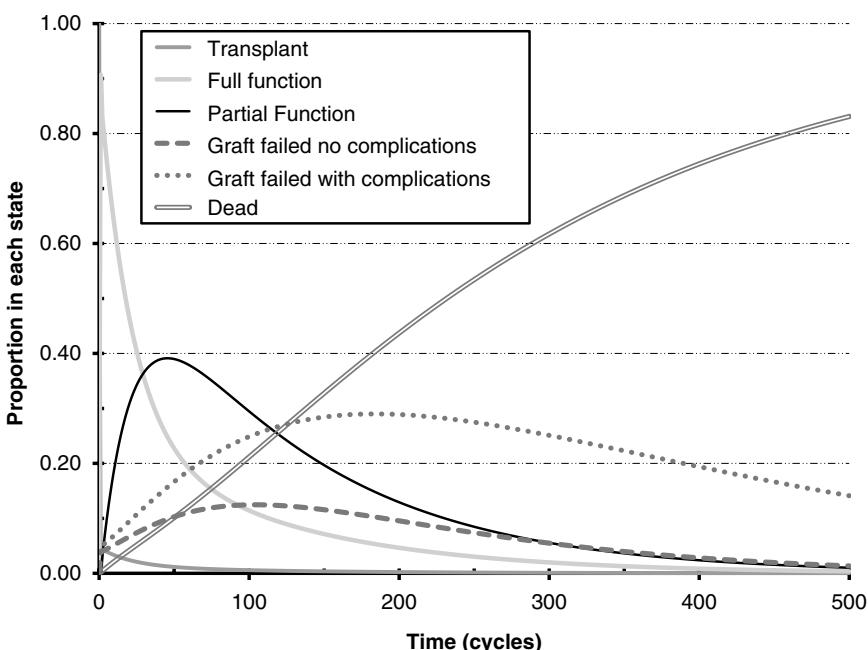
Graphical depictions of a Markov trace get complex very quickly. Figure 5.4 shows a Markov trace from a model with 6 states in 1 treatment arm over 500 cycles.

**Table 5.3** Markov trace for a simple Markov model

	Well	Disease	Dead
$T=0$	100	0	0
$T=1$	70	20	10
$T=2$	51	26	23
$T=3$	38.9	25.3	35.9
$T=4$	29.39	23.14	47.47
$T=5$	22.887	19.762	57.351



**Fig. 5.3** Graphical depiction of the Markov trace for a simple Markov model



**Fig. 5.4** Graphical depiction of a Markov trace for a more complex Markov model

## 5.6 Cycle Length, Time Horizon and Discounting

Whilst we have seen how people transition between states over time, how do we decide how long each period or cycle is? The length of the cycle is influenced by a number of different factors. The cycle length is the minimum time people spend in a state (all members of the cohort will spend at least one cycle in the state they begin the model in). Thus the cycle length may be determined as the minimum period of time that people spend in a state. This is often informed by the smallest clinically meaningful item included as a distinct event. Sometimes we combine events, for example, events within a hospitalisation may be grouped together rather than considered separately. Cycle length normally doesn't matter much analytically, but it can be important to the amount of work involved. It is easier to model a series of small, well-defined cycles than work out what happens in one long (and potentially messy) cycle. In general, finding costs and benefits to assign to health states is not that much harder for short or long cycle lengths.

In our previous example we saw how the Markov trace uses the transition probabilities to calculate the movement between groups. However, transition probabilities depend in part on the cycle length. We can use a simple Markov model to demonstrate this (Fig. 5.5).

This time our Markov model has only two health states, Well and Dead. If the chance of dying in 1 year is 25 % and we have a corresponding cycle length of 1 year, then the probability of moving from Well to Dead is  $P(\text{Dead}_{+1} | \text{Well}_t) = 0.25$ . But if the cycle length is only, for example, 6 months, then the chance of dying in cycle 1 is not 0.125 as one might expect. This may seem counterintuitive, but probabilities cannot be divided over time; instead we have to use rates which can be divided this way. We need to convert the 1-year probability to a 1-year rate, divide this by two to get a 6-month rate and then convert this to a 6-month probability. Assuming constant hazards, the equations we need are

$$r = \ln(1 - p)$$

and

$$p = 1 - e^{-r}$$

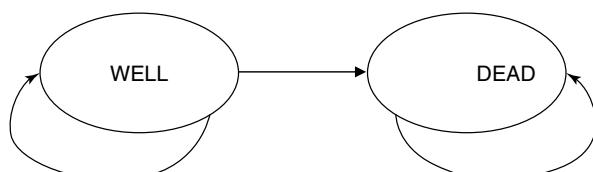
where  $r$  is the rate and  $p$  is the probability.

In this example, the 12-month probability is 0.25:

The 12 month rate is  $\ln(1 - 0.25) = 0.288$

The 6 month rate is  $0.288 \div 2 = 0.144$

The 6 month probability is  $1 - e^{-0.143841} = 0.134$



**Fig. 5.5** Simple Markov model

**Table 5.4** Markov trace for our simple Markov model with more than one starting state

	Well	Disease	Dead
T=0	60	40	0
T=1	46	36	18
T=2	35.8	30.8	33.4
T=3	28.14	25.64	46.22

Try this calculation for yourself using the exercise below. The answer is shown at the bottom of the page.<sup>1</sup>

**Exercise 5.1** Assuming constant hazards, find the per cycle probability if we have a cycle length of 1 month and the 12-month probability is 50 %.

Associated with cycle length is the period that we follow this process for. The time horizon is the total period of time over which the models runs. This will depend on when the relevant costs or benefits of the model stop happening, as it should be long enough to capture meaningful differences in costs and outcomes between the intervention and comparator (CADTH 2006; NICE 2013). In general a lifetime horizon is thought to be the default position and will be best for most chronic conditions. Some types of question will have a shorter horizon for a variety of reasons: it might be because costs and benefits stop accruing, because this is in line with decision makers' requirements or that there is such limited evidence for the longer term that the decision maker chooses a shorter time horizon. In the latter case, it is important to note that this must be the decision maker's judgement, not the analyst's. There is no natural end to a Markov tree. Consider our simple example in Sect. 5.5. Over time, the proportion of people in the Well and Disease health states will fall, but they will never reach zero – although they will of course become exceedingly small! Markov models need to have a stopping point. In taking a lifetime horizon, models are often limited to those in the cohort reaching 100 years of age (i.e. the cohort excludes those living beyond 100 years of age) – although this is arbitrary.

In respect of where individuals in the cohort start within the model, normally models will identify one state as the 'starting state'. The starting state will depend on the decision problem. Consider, for example, a model for an intervention to prevent hospital-acquired infection. The starting state for the population in this model might be 'free from infection'. Occasionally a state might be invented just to represent the starting state, or there may be more than one starting state. Consider our previous model which consisted of three health states: Well, Disease and Dead. We had previously assumed that all those entering the model were in the Well state, but in fact some individuals may already have the disease of interest. In which case our Markov trace could look like Table 5.4.

An apparently small but occasionally important consideration with Markov models is when during a cycle do we believe events occur? It doesn't make any sense to assume that they all occur at the beginning or end of the cycle. It is more

---

<sup>1</sup> Answer 0.0561

**Table 5.5** Discount factor

Time (years)	Calculation	\$
0	$1/(1+0.05)^0$	1
1	$1/(1+0.05)^1$	0.9524
2	$1/(1+0.05)^2$	0.9070
3	$1/(1+0.05)^3$	0.8638
4	$1/(1+0.05)^4$	0.8227

plausible that they are evenly or at least symmetrically distributed over the entire cycle period. If you accept this line of reasoning, then it follows that on average, the events will occur approximately halfway through the cycle. However in discrete state models, such as Markov state models, all the patients in each state accrue the full costs and health for each cycle. This means that the Markov trace for costs and outcomes will overstate the underlying continuous processes that we are trying to model. A simple and easily implemented solution to this problem is to shift the allocation of patients by half a cycle, so that the costs and outcomes attributed to each state in each cycle are based on half the state membership from the current cycle and half the state membership from the next cycle. To implement this correction is quite simple. In calculating the *total* costs and outcomes, we replace the first period result with the total of 50 % of the first period result and 50 % of the final period result. Naimark et al. (2008) provide an excellent and accessible exploration of the thinking behind the half cycle correction and how it has the desired effect.

Finally when considering the time horizon, we also we need to consider applying a discount factor given that the value of costs and benefits depends on both their *value* and *when* they occur. A consistent *value* for costs is obtained by using costs adjusted to refer to the same currency/year. In respect of *when* they occur, the value of future costs and benefits decreases if they occur further into the future. We can use the following formula to obtain a discount factor

$$1/(1+r)^n$$

where  $r$  is the discount rate and  $n$  is the number of years from now. An example is shown in Table 5.5 that assumes  $r=5\%$ .

Choice of the discount rate may be dictated by national guidelines. For example, in Canada, the Canadian Agency for Drugs and Technologies in Health (CADTH) guideline recommends that a rate of 5 % per annum is used for costs and health effects (CADTH 2006), whereas in the UK, NICE recommends a rate of 3.5 % per annum be applied to both. However, NICE goes on to suggest that over long periods ( $>30$  years) a sensitivity analysis might be considered using a discount rate of 1.5 % for both costs and health effects (NICE 2013; O'Mahoney and Paulden 2014). Similarly CADTH recommends sensitivity analysis using 0 % to show the impact of discounting (CADTH 2006). In the absence of country-specific guidance, the convention is to use a rate consistent with existing literature, which allows comparisons to be made between different studies. There has been a substantial debate as to the principles that should guide the choice of discount rate and whether

costs and outcomes should be discounted at the same rate. This is beyond the scope of this book, but interested readers are recommended to review Claxton and colleagues (2011).

## 5.7 Summary

- Markov models are most useful when health events repeat over time or have longer-term health effects; when the effect of treatment either stops quickly after an initial treatment or continues at its earlier level; and when the risk of different health events does not depend on the patient's prior history.
- The clinical pathway is fundamental given Markov models are structured around health states and movements between them. Influence diagrams are a great way to visually represent your model.
- Transition probabilities predict how people will move from one health state to another.
- The Markov trace captures the numbers or proportion of people in each health state over time. The Markov trace uses the transition probabilities to calculate the movement between groups within your model.
- The cycle length is the minimum period of time that people spend in a health state. This is often informed by the smallest clinically meaningful item included as a distinct event.
- The time horizon is the total period of time we follow the cycles and should be chosen to reflect when relevant costs or benefits of the model stop happening.
- For models following a cohort for >1 year, a discount factor should be applied.

## References

- Canadian Agency for Drugs and Technologies in Health (2006) Guidelines for the economic evaluation of health technologies: Canada, 3rd edn. Canadian Agency for Drugs and Technologies in Health, Ottawa
- Claxton K, Paulden M, Gravelle H, Brouwer W, Culyer AJ (2011) Discounting and decision making in the economic evaluation of health-care technologies. *Health Econ* 20(1):2–15. doi:[10.1002/hec.1612](https://doi.org/10.1002/hec.1612), Epub 2010 May 12
- Evans KW, Boan JA, Evans JL, Shuaib A (1997) Economic evaluation of oral sumatriptan compared with oral caffeine/ergotamine for migraine. *Pharmacoeconomics* 12(5):565–577
- Kielhorn A, von der Schulenburg JMG (2000) The health economics handbook. Adis International, Chester
- Naimark DMJ, Bott M, Krahn M (2008) The half-cycle correction explained: two alternative pedagogical approaches. *Med Decis Making* 28(5):706–712
- National Institute for Health and Care Excellence (2013) Guide to the methods of technology appraisal 2013. <http://publications.nice.org.uk/pmg9>
- O'Mahony JF, Paulden M (2014) NICE's selective application of differential discounting: ambiguous, inconsistent, and unjustified. *Value Health* 17:493–496, <http://dx.doi.org/10.1016/j.jval.2013.02.014>. Published online ahead of print May 17 2013

# Chapter 6

## Probability Distributions for Effectiveness Parameters

**Abstract** When we construct a decision model for cost effectiveness analysis, we are representing the flow of patients through a clinical pathway. As patients flow through a pathway, represented by health states in our model, they will experience clinical events. Clinical events are measured in real life by conducting clinical research or audit and are reported as health outcomes. It is these health outcomes that are represented in our model by effectiveness parameters. In this chapter we consider how these parameters can be incorporated into a model in a manner that captures the uncertainty in estimated effectiveness.

### 6.1 Introduction

At the heart of any cost effectiveness model are the effectiveness parameters for the technologies being compared. In this chapter we consider how these parameters can be incorporated into a model in a manner that captures the uncertainty in estimated effectiveness. The material covered in this chapter represents a solid foundation for students seeking to undertake work in this setting. There are more advanced methods available that are beyond the remit of this text, especially with regard to effectiveness evidence derived from advanced evidence synthesis methods such as network meta-analysis. Researchers wishing to utilise these types of data are encouraged to undertake more advanced study and/or collaborate with peers with the appropriate advanced training.

In this chapter we consider effectiveness data, which is characterised as cross-sectional risk and time-to-event risk data. We introduce a range of distributions for characterising probabilities, relative risks and time-to-event effectiveness information. Sections 6.2 and 6.3 provide background information on what we mean by effectiveness data and how to fit a distribution. Sections 6.4, 6.5, 6.6 and 6.7 then consider specific types of data and distribution combinations. Section 6.8 briefly

considers the issue of correlation between parameters. The chapter concludes with a summary of the key points.

## 6.2 What Do We Mean by Effectiveness Parameters?

When we construct a decision model for cost effectiveness analysis, we are representing the flow of patients through a clinical pathway. As patients flow through a pathway, represented by health states in our model, they will experience clinical events. Clinical events are measured in real life by conducting clinical research or audit and are reported as health outcomes. It is these health outcomes that are represented in our model by effectiveness parameters.

We can consider health outcomes for a single group of patients. Here, we are representing activity as the rate of events over time or, in the case of a discrete-time Markov model, the probability of an event within a fixed period of time. It is these probability parameters that comprise the transition probabilities within our transition matrix.

We might also wish to consider the differences in health outcomes between two or more groups of patients. In this case we could simply consider the different event probabilities in each group in isolation. It is often more useful or more analytically powerful to represent a comparison between groups as a specific parameter. In this case we need to consider relative effectiveness as opposed to absolute effectiveness. We need to consider relative risk and related concepts such as the odds ratio or hazard ratio, all of which will be considered in this chapter.

In previous chapters you learnt about the importance of conducting a probabilistic sensitivity analysis (PSA). Such analysis conducted using Monte Carlo simulation requires a random draw from a mathematically defined probability distribution with a mean and variance. The remainder of this chapter describes how to represent effectiveness parameters.

### 6.2.1 *Obtaining Information on Effectiveness*

In an ideal world we would always directly observe any health outcomes of interest in a relevant patient population. The forthcoming data can then be used to directly inform relevant parameters in our model. In a PSA this can be achieved either by sampling directly from the data or by sampling from a mathematical distribution fitted very closely to the data using statistical software.

Very frequently, when conducting a model-based economic evaluation, a primary dataset is not available to inform all the parameters. In this case the parameter distributions must be defined through other means. Most commonly information can be sought from the published literature (see Chap. 2). Sources include published clinical studies in isolation or multiple studies that have been identified through

systematic review and, where appropriate, meta-analysis. Methods for evidence synthesis to inform model parameters are beyond the scope of this book. For the interested reader, Welton et al. (2014) provide excellent coverage of this area.

When seeking information from the published literature to inform model parameterisation, the modeller is often faced with incomplete or inadequate information to fit mathematical distributions. Despite this it is often possible to estimate distributions provided some key elements are included. The task should become easier as reporting standards for scientific publications are increasingly adopted by high-quality journals.

### 6.3 Choosing Distributions for Effectiveness Parameters

The first step in assigning a distribution to a parameter is to choose the appropriate distribution. An important consideration when doing this is achieving the correct balance between accuracy of fit to the data and simplicity. Accuracy of fit will improve model precision, whereas simplicity will protect against user error and facilitate peer review, reproducibility and computational feasibility. In theory there are an infinite number of mathematical distributions to choose from. An individual distribution is described using one or more distribution parameters (as opposed to what we have been referring to as model parameters). The more parameters a distribution has, the more flexible will be that distribution and the better potential fit will be achieved.



Each distribution has its own unique characteristics. The challenge of the modeller is to match those characteristics with the properties of a model parameter. The simple and familiar option of the Normal distribution is an option justified by the central limit theorem. The use of a Normal distribution can, however, produce

unsuitable values for certain model parameters during Monte Carlo sampling. A good example is a probability parameter. Probabilities are bound by zero and one. A distribution that encompasses values outside this range may produce implausible values during Monte Carlo sampling. Similarly, a relative risk should have a lower bound of zero. For both these types of parameters, the Normal distribution may produce values outside the plausible range, particularly where the variance is large or the mean of the distribution is close to the upper or lower bound. Whilst there are many parameters, many decision analytic cost effectiveness models make use of quite a small subset: see Box 6.1.

**Box 6.1: Distributions Commonly Used in Decision Analytic Cost Effectiveness Models**

Distribution	Parameters	Application
Normal	Mean ( $\mu$ ); standard deviation ( $\sigma$ )	Effectiveness, utilities
Beta	$\alpha, \beta$	Effectiveness, utilities
Dirichlet	$\alpha, \beta$	Effectiveness
Gamma	$\alpha, \beta$	Costs, utilities
LogNormal	Mean ( $\mu$ ); standard deviation ( $\sigma$ )	Costs, effectiveness
Weibull	$\lambda, \kappa$	Effectiveness
Gompertz	$Y, \lambda$	Effectiveness
Exponential	$\lambda$	Effectiveness

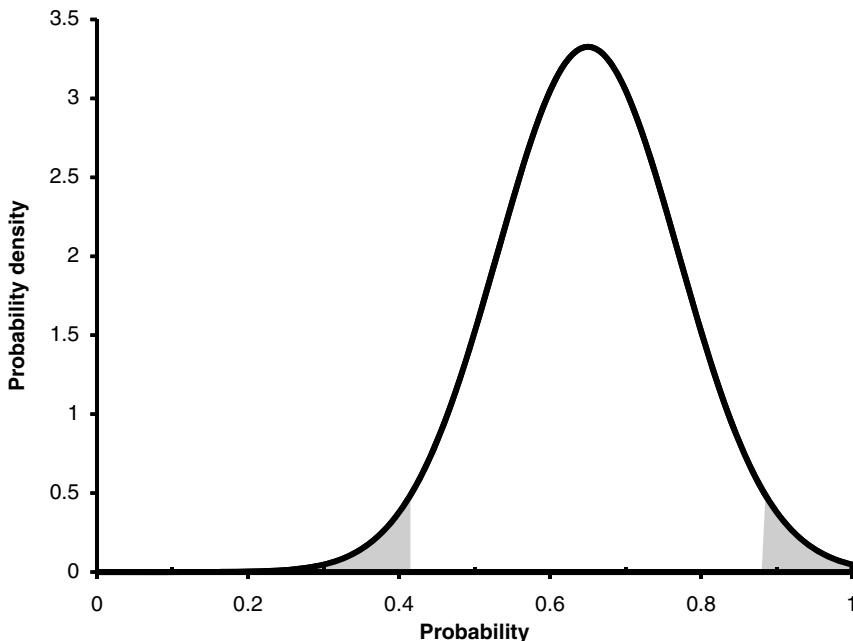
### 6.3.1 Fitting a Distribution

Several techniques can be used to fit distributions. Let us take a simple example to illustrate how a Normal distribution can be fitted to a probability parameter. It is common for publications to report a mean with 95 % confidence intervals. In order to specify a distribution in Excel or many other software packages, we need the mean and standard deviation of the distribution. Let us take an example where the mean of a probability is 0.65, and we are told that the 95 % confidence interval is 0.415–0.885. We know from the probability density function of the Normal distribution that the 95 % confidence interval covers 3.92 standard deviations ( $2 \times 1.96$ ). Therefore the standard deviation equals  $(0.885 - 0.415) / 3.92 = 0.12$ .

To check you understand this, calculate the mean and standard deviation for a normally distributed parameter that has a 95 % confidence interval of 1.514–2.423. The answer is at the bottom of the page<sup>1</sup>.

The Normal distribution (Fig. 6.1) is defined by two parameters: mean/median/ $\text{mode} = \mu$  and variance =  $\sigma^2$ . These parameters are referred to as the moments of the distribution. Any distribution can be defined from its moments, and it is these that

<sup>1</sup>Answer:  $2.423 - 1.514 = 0.909$ ;  $\text{SE} = 0.909 / 3.92 = 0.232$ ;  $\text{mean} = 1.514 + (1.96 \times 0.232) = 1.969$ .



**Fig. 6.1** The Normal distribution

need to be calculated from available information in order to fit a distribution. Applying this technique is referred to as the method of moments.

## 6.4 Beta Distribution for Probabilities

Although the central limit theorem offers the Normal distribution as a candidate distribution for any parameter, this can lead to problems in the context of a model where the properties of the Normal distribution are not in line with the properties of a specific parameter. A probability parameter is a good example. A probability is constrained to lie between zero and one. It can be seen in Fig. 6.2 that the Normal distribution, fitted to information about our example probability, will include values above one at the extreme of its range. Monte Carlo simulation from this distribution will produce rare samples above one. Although this could be dealt with by truncating the distribution, a more coherent solution would be to choose an alternative distribution with characteristics matching those of a probability.

The Beta distribution offers a good solution. The Beta distribution is a unimodal distribution constrained to lie between zero and one, thus meeting our requirements for a probability parameter. It has two parameters referred to as alpha and beta ( $\alpha$  and  $\beta$ ). Some example Beta distributions are shown in Fig. 6.3.

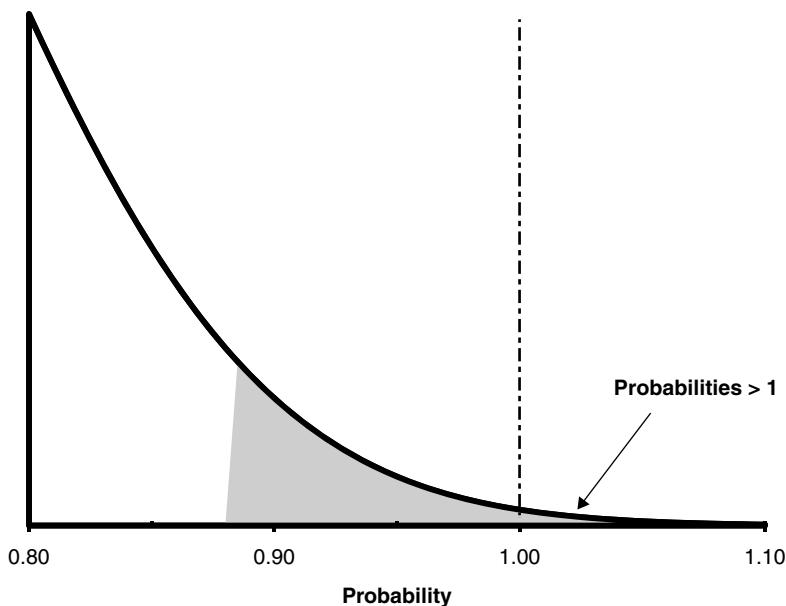


Fig. 6.2 Normal distribution probability exceeds 1.0

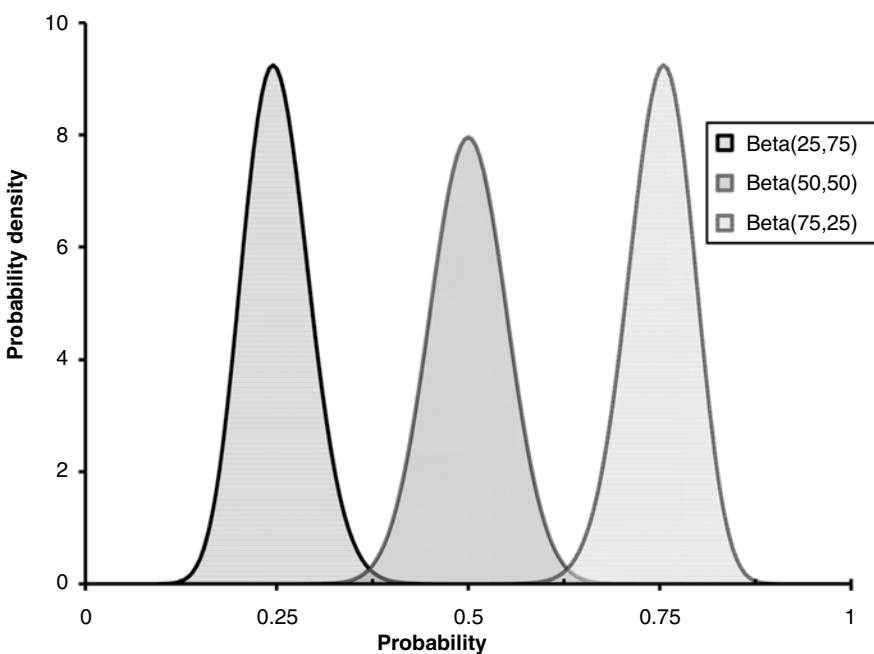


Fig. 6.3 Example Beta distributions

It is particularly easy to fit a Beta distribution where the numbers of observed events out of a known population size are reported. For example, if 5 patients die and 1,426 remain alive out of a total population of 1,431 patients, then the Beta distribution representing the probability of death can simply be defined as  $\alpha = 5$  and  $\beta = 1,426$ .

However, we frequently do not have this level of information. Under these circumstances, we can work out  $\alpha$  and  $\beta$ , if we have the mean and the variance. If  $x$  is distributed Beta  $[\alpha, \beta]$ , we know that

$$x \text{ has an expected value } E[x] = \frac{\alpha}{\alpha + \beta}$$

$$x \text{ has variance } V[x] = \frac{\alpha\beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)}$$

The parameters  $\alpha$  and  $\beta$  can be calculated from the mean and variance by first defining

$$\alpha + \beta = \frac{E[x](1 - E[x])}{s^2} - 1$$

where  $s$  is the sample standard deviation. Then

$$\alpha = E[x](\alpha + \beta)$$

and

$$\beta = (\alpha + \beta) - \alpha$$

For our example, we can therefore calculate the parameters as

$$\alpha = E[x](\alpha + \beta) = 0.6514.80 = 9.62$$

and

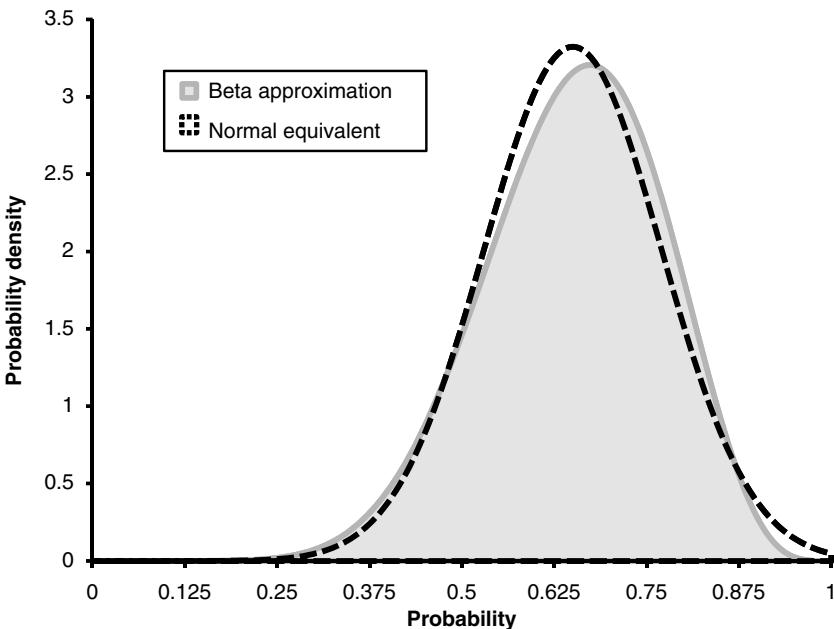
$$\beta = (\alpha + \beta) - \alpha = 14.80 - 9.62 = 5.18$$

To check you follow this, calculate  $\alpha$  and  $\beta$  when  $E[x] = 0.40$  and  $V[x] = 0.04$ . The answer is at the bottom of the page.<sup>2</sup>

Figure 6.4 shows how the Normal distribution and Beta distribution differ for our example data. The Normal distribution exceeds 1.0, and the Beta distribution is less symmetrical. Whilst the differences may appear small, the use of the wrong distribution could impact on a reimbursement decision in a number of ways. Most obvi-

---

<sup>2</sup>Answer to Beta distribution practice:  $0.4 \times 0.6 = 0.24$ ;  $0.24/0.04 = 6$ .  $\alpha + \beta = 5$ ,  $\alpha = 2$ ,  $\beta = 3$ .



**Fig. 6.4** Beta distribution with  $E[x]=0.65$  and  $s=0.12$  with equivalent Normal distribution superimposed

ously, the use of a model that included impossible values – such as probabilities exceeding 1 – would be open to judicial challenge by patients and manufacturers who disagreed with the decision. Even if the error goes unnoticed, any correlation between model parameters means that this can have a ‘multiplier effect’ in terms of their impact on the estimates of costs and outcomes in probabilistic analyses. When the differences in the value of the alternative technologies are small, as is often the case, such errors may actually change the decision.

## 6.5 Dirichlet Distribution for Multinomial Probabilities

Whilst the Beta distribution can be used for a probability parameter, it is only useful for binomial probabilities. We will often need to include a multinomial parameter, for example, where we wish to represent the probability of transitioning from a health state in a Markov model into two or more alternative states. The Dirichlet distribution is the multivariate generalisation of the Beta distribution. It is easy to fit where the counts from which probabilities were estimated are available; like the Beta distribution, the parameters take the values of the observed events.

Many software packages, including Excel, lack a Dirichlet function. One solution is to use sequential Beta distributions, but this is artificial and lacks computational

efficiency. A preferred method of implementation is to use independent single parameter Gamma distributions. The Gamma distribution is described in detail in the section on distributions for cost parameters. In Excel, the Gamma distribution has two parameters (alpha and beta) [a random draw is taken using the function =Gamma.Inv(Rand(),  $\alpha, \beta$ )]. To implement a draw from a Dirichlet distribution, the beta parameter is set to one, and the alpha parameter takes the value for the number of observed events. Each of the multinomial probabilities is then derived by dividing draw for the outcome probability of interest by the sum of the draws for all the outcomes in the distribution.

As an example, consider the transition matrix for the exemplar Markov model from Chapter 5 (Tables 5.1 and 5.2). The probability of moving from the state Well to Disease or Dead or staying in Well can be represented by a Dirichlet distribution dirich ( $u, v, w$ ). If the probabilities were estimated from an observed sample of 100 events, where at the end of time period 1, 70 patients remained well, 20 developed disease and 10 had died, then the parameter is defined as dirich (70, 20, 10). In Excel this can be implemented using independent Gamma distributions:

$$u \text{ from Gamma}(70, 1) \text{ then } P(\text{Well}_{t+1} | \text{Well}_t) = u/(u+v+w)$$

$$v \text{ from Gamma}(20, 1) \text{ then } P(\text{Disease}_{t+1} | \text{Well}_t) = v/(u+v+w)$$

$$w \text{ from Gamma}(10, 1) \text{ then } P(\text{Dead}_{t+1} | \text{Well}_t) = w/(u+v+w)$$

## 6.6 Normal Distribution for Log-Relative Risk

Having developed a model representing standard care without an intervention of interest, a second version of the model is necessary that includes the intervention and its effect. The two versions of the model can then be compared to provide an incremental analysis. Let us again consider the three-state Markov model from Chap. 5. If our new intervention of interest reduces the probability of moving from the Well state to the Disease state, then its effect will be to reduce  $P(\text{Disease}_{t+1} | \text{Well}_t)$ . One way to do this might be to simply define a separate probability. In many situations, however, it will be better to keep the baseline  $P(\text{Disease}_{t+1} | \text{Well}_t)$  identical in both versions of the model but apply a relative risk in the version that contains the new intervention. The version of the transition matrix for the intervention model is shown in Table 6.1.

It is mathematically coherent to include the relative risk in a model as log-relative risk, which is normally distributed. After taking a random draw from the log-relative risk, the result is exponentiated before multiplying with the baseline probability.

**Table 6.1** Transition probabilities for our simple Markov model with the treatment effect of an intervention applied

		End of period $t$ (start of period $t+1$ )		
		Well	Disease	Dead
Start of period $t$	<b>Well</b>	$1 - P(\text{Disease}_{t+1}   \text{Well}_t) \times RR - P(\text{Dead}_{t+1}   \text{Well}_t)$	$P(\text{Disease}_{t+1}   \text{Well}_t) \times RR$	$P(\text{Dead}_{t+1}   \text{Well}_t)$
	<b>Disease</b>	$P(\text{Well}_{t+1}   \text{Disease}_t)$	$P(\text{Disease}_{t+1}   \text{Disease}_t)$	$P(\text{Dead}_{t+1}   \text{Disease}_t)$
	<b>Dead</b>	$P(\text{Well}_{t+1}   \text{Dead}_t)$	$P(\text{Disease}_{t+1}   \text{Dead}_t)$	$P(\text{Dead}_{t+1}   \text{Dead}_t)$

If the effect size is represented by a linear predictor derived from a regression or multiparameter model, it will be necessary to take account of the correlation between these parameters (see Chap. 8).

Often an effectiveness estimate will be reported as an odds ratio. This is common where logistic regression is used to adjust for covariates. In order to apply the effect size within our transition matrix, it will be necessary to convert the odds ratio into a relative risk. This is possible where the baseline event probability ( $p_0$ ) is known:

$$RR = \frac{OR}{1 - p_0 + (p_0 \times OR)}$$

where RR=relative risk, OR=odds ratio and  $p_0$ =baseline event probability.

## 6.7 Survival Analysis for Time-to-Event Data

The transition probabilities in a Markov model represent the probability of an event occurring during a period of time, the duration of which is defined by the model cycle length. Events can be death or any other event that results in a transition between health states in our model or incurs costs. The probability is a discrete representation of the risk or ‘hazard’ of an event occurring at any of the specific points in time during that cycle. The risk may be constant or may rise or fall over time. If the underlying risk changes over time, then consecutive transition probabilities will be different and cannot simply be represented by a stand-alone probability parameter.

If we want to describe the changing risk in our Markov model – correctly referred to in the case of nonconstant transition probability as a semi-Markov or modified Markov model – then it is necessary to do so using a numeric function, referred to as a hazard function,  $h(t)$ . As time passes and risk accumulates, we are able to define the cumulative hazard function,  $H(t)$ :

$$H(t) = \int_0^t h(u) du$$

Derived from the cumulative hazard function is the survival function which describes the occurrence of events over time and is responsible for producing the familiar survival curves we see in the scientific literature:

$$S(t) = e^{-H(t)}$$

The transition probabilities that we need for our Markov model are derived from the survival function. A transition probability, or the probability of an event occurring over a model cycle, is the (proportional) drop in the survival function over that defined time interval:

$$\begin{aligned} p(t) &= 1 - \frac{S(t)}{S(t-1)} \\ &= 1 - e^{H(t-1)-H(t)} \end{aligned}$$

It is relevant to note that the functions described here can be defined using closed formulas (parametric) or can be defined directly from observed data using non-parametric methods such as the Kaplan-Meier or product limit method. Whilst survival data can be incorporated directly in a model, the focus of this chapter is on the use of parametric methods. When the analyst has access to raw data, the mean and standard deviation of the distribution can be derived by fitting the exponential distribution using a standard statistical package Latimer (2011). When estimating lambda from a published analysis without access to patient-level data, computer digitisation programmes can be used to reproduce the survival data.

### ***6.7.1 The Exponential Distribution***

The exponential (or correctly, negative exponential) distribution assumes a constant hazard rate over time. This means that transition probabilities will be identical between cycles. Its use has several advantages over a simple probability parameter, in particular where adjustment for covariates is required.

The survivor function of the exponential distribution is

$$S(t) = e^{-\lambda t}$$

The underlying hazard function is

$$h(t) = \lambda$$

where  $\lambda > 0$  is the single model coefficient.

In PSA, lambda parameter values are sampled from a LogNormal distribution. If patient-level data cannot be obtained or derived from the published literature, the lambda parameter can be back calculated from the median survival and confidence intervals.

Transition probabilities for a standard care Markov model are calculated from lambda by

$$tp_0 = 1 - e^{-\lambda u}$$

where  $u$  is the cycle length.

The transition probability for an intervention arm can either be estimated as a separate independent lambda parameter if appropriate or, more usually, will require the underlying hazard (rather than the transition probability) to be multiplied by a hazard ratio for each cycle:

$$tp_1 = 1 - e^{-\eta \lambda u}$$

where  $\eta$  is the hazard ratio.

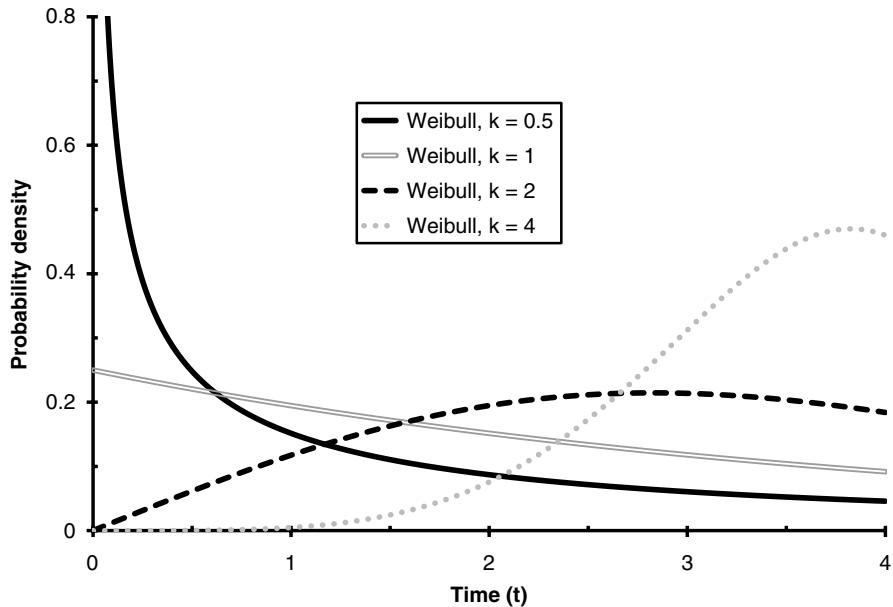
A significant drawback of the exponential distribution is that the hazard remains constant for an indefinite length of time. An arbitrary time point can be specified at which this becomes zero or changes, but this must be justified and tested in a deterministic sensitivity analysis. The piecemeal exponential distribution offers a worthwhile and underused solution where the hazard rate is defined as different over different time periods. Here, lambda is estimated separately for each time period. Each time period can potentially have a fitted survival model with an estimate of uncertainty and covariates. However, this method will inevitably be criticised for the unavoidable subjectivity of the choice of time periods. Hence it will be important to test the impact of using alternative time periods on the results.

### 6.7.2 The Weibull Distribution

The Weibull distribution is more flexible than the exponential. As well as a scale parameter  $\lambda > 0$ , it has an additional shape parameter  $k > 0$  that allows increasing, constant or decreasing hazards. The bigger is  $\lambda$ , the quicker the survival function falls.  $k$  defines how the hazard changes over time. When  $k=1$  the risk is constant (equivalent to the exponential function). When  $k > 1$  the risk increases over time. When  $k < 1$  it decreases over time. The effect of varying the shape parameter can be seen in Fig. 6.5.

The survivor function of the Weibull distribution is

$$S(t) = e^{-(\lambda t)^k}$$



**Fig. 6.5** Weibull hazard functions with different shape parameters

The hazard function is

$$h(t) = \lambda k (\lambda t)^{k-1}$$

### 6.7.3 The Gompertz Distribution

Although originally introduced by Gompertz in 1825 as a model for human mortality, it was in the 1960s that A.K. Laird for the first time successfully used the Gompertz curve to fit data of growth of tumours (Laird 1964). Like the Weibull distribution, the Gompertz distribution has two parameters (shape and scale), and its hazard increases or decreases monotonically with an inherent proportional hazards assumption.

The survivor function is

$$S(t) = \exp \left\{ -\frac{\lambda}{\gamma} (1 - e^{\gamma t}) \right\}$$

The hazard function is

$$h(t) = \lambda e^{\gamma t}$$

The cumulative hazard is

$$H(t) = \frac{-\lambda}{\gamma} (1 - e^{\gamma t})$$

Transition probabilities for each model cycle at time  $t$  with cycle length  $u$  are given by

$$tp(t, u) = 1 - \exp \left\{ \frac{\lambda}{\gamma} (e^{\gamma(t-u)} - e^{\gamma t}) \right\}$$

For the intervention arm, where  $\eta$  is the hazard ratio, transition probabilities are given by

$$tp(t, u) = 1 - \exp \left\{ \frac{\eta \lambda}{\gamma} (e^{\gamma(t-u)} - e^{\gamma t}) \right\}$$

#### 6.7.4 Choice of Distribution for Time-to-Event Data

The basic principle for choosing which distribution to use is quite straightforward – choose the distribution that fits the data best. The easiest way to see what this means in practice is to examine an example. In Figs. 6.6 and 6.7 we explore the alternative models for distant disease-free survival in breast cancer patients. The models were fitted to data from a 15-year follow-up study reported by Dent et al (2007). The candidate distributions include the Exponential, Weibull, LogNormal, Dent et al (2007) and Gompertz distributions. Contained within Fig. 6.6 is the Kaplan-Meier plot from which the data were derived using a digitisation technique (Guyot et al 2012). Looking at the plots in Fig. 6.6 we can see that the Weibull, Exponential and LogNormal all substantially overestimate survival in the early years and then underestimate survival in the later years. By comparison, the Gompertz function matches the Kaplan-Meier curves remarkably closely. There is a moderate underestimation of survival for the first 3 years, but from 3 years onwards the Gompertz curve maps the observed events remarkably well. On this basis, we could defend choosing the Gompertz function. In the absence of the Gompertz function, the exponential is clearly the worst of the remaining options, whilst the LogNormal is similarly poor. This demonstrates the importance of a full range of alternative functions. Figure 6.6 illustrates the substantial differences in the hazard functions associated with each model (Fig. 6.7).

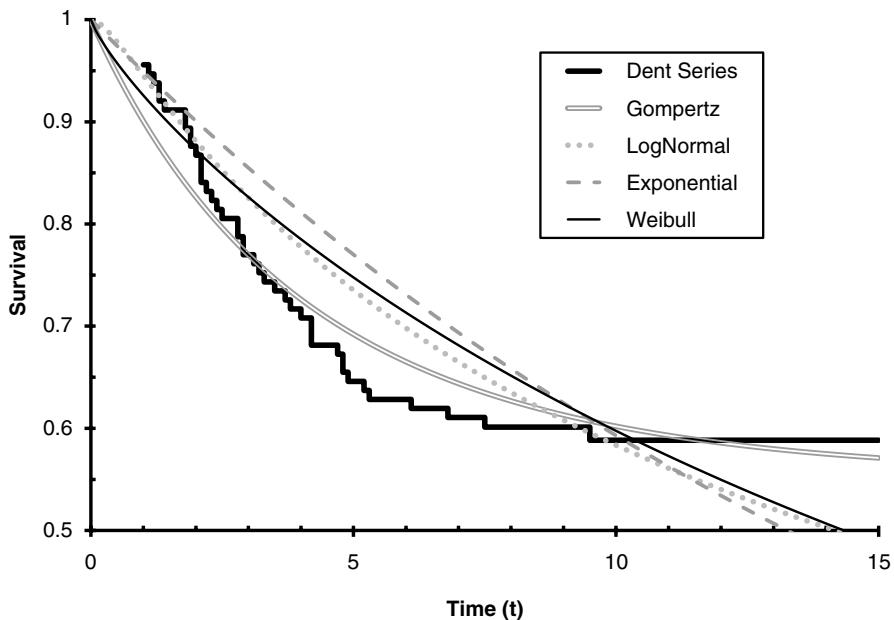


Fig. 6.6 Kaplan-Meier plot with fitted parametric survival distributions

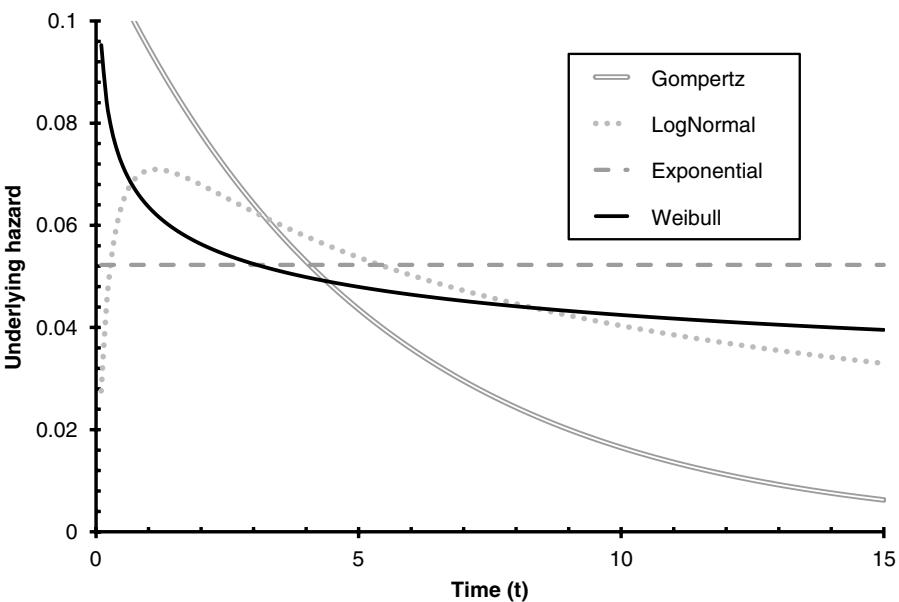
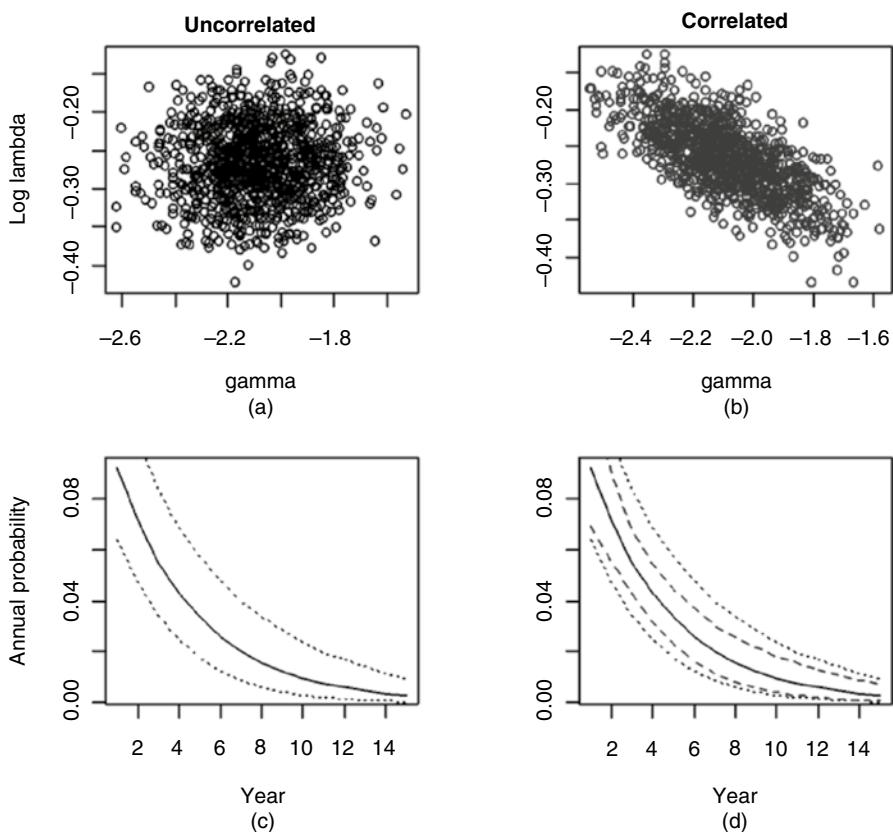


Fig. 6.7 Underlying hazard for Kaplan-Meier plot and fitted distribution (Dent 2007)

## 6.8 Parameter Correlation in Survival Analysis

It is worth noting that when the chosen parametric survival distribution is defined by more than one parameter, the correlation between these two parameters needs to be considered, within the PSA. PSA using Monte Carlo simulation requires a random draw from the parameters. It is highly likely that the parameters of a survival distribution will be correlated. Ignoring this correlation will overestimate uncertainty in the model as seen in Fig. 6.8. (The dotted lines represent the interval assuming uncorrelated parameters, with the dashed lines showing the (narrower) intervals for correlated parameters). In order to represent parameter correlation in our model, it will be necessary to use the variance-covariance matrix produced by the regression used to fit the survival model. Most statistical packages will provide this. Chapter 8 describes how to use the variance-covariance matrix to construct a random draw that takes account of the correlation between the parameters.



**Fig. 6.8** Correlations in survival function data

The correlation between the gamma and lambda parameters of the Gompertz survival distribution reduces the width of the 95 % confidence intervals (dotted and dashed lines) around the annual probability of recurrence.

## 6.9 Summary

- Choosing an appropriate distribution for characterising effectiveness parameters is essential to providing robust cost effectiveness evidence to support resource allocation decisions.
- Ideally, the analyst will have patient-level data to do this from – however, this is frequently not the case.
- There are substantial downsides to using a Normal distribution to characterise an absolute risk parameter.
- For categorical effectiveness data the Beta or Dirichlet distribution is often appropriate.
- For time-to-event effectiveness data, exponential, Weibull and Gompertz distributions are frequently used. Alternatives include the LogNormal and several other more complex distributions. The choice of distribution should be based upon how well the alternatives fit the existing data.
- Where effectiveness data are characterised by two or more parameters, consideration of the correlation between these parameters will be required to avoid the misspecification of the uncertainty – see Chap. 8 for details on how to do this.

## References

- Dent R et al (2007) Triple-negative breast cancer: clinical features and patterns of recurrence. *Clin Cancer Res* 13:4429
- Guyot P et al (2012) Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med Res Methodol* 12:19. doi: [10.1186/1471-2288-12-9](https://doi.org/10.1186/1471-2288-12-9)
- Laird AK (1964) Dynamics of tumor growth. *Br J Cancer* 18(3):490–502
- Latimer N (2011) NICE DSU technical support document 14: survival analysis for economic evaluations alongside clinical trials – extrapolations with patient level data. Report by the DSU. <http://www.nicedsu.org.uk/NICE%20DSU%20TSD%20Survival%20analysis.updated%20March%202013.v2.pdf>. Accessed 11 Oct 2014
- Welton N, Sutton AJ, Cooper NJ, Abrams KR, Ades AE (2014) Evidence synthesis for decision making in health care. Statistics in practice. Wiley, Chichester England

# Chapter 7

## Probability Distributions for Cost and Utility Parameters

**Abstract** There are often greater uncertainties affecting the cost estimates in an analysis than other parameters. As a result, the appropriate characterisation of the uncertainty in cost parameters is important to provide decision makers with robust evidence that can withstand careful appraisal. Similarly, the utility parameter distribution should reflect the uncertainty around the expected value for the utility of the model health state for a homogenous patient population. Having considered the range of distributions that are most frequently used for characterising effectiveness parameters and the uncertainty regarding their true values in Chap. 6, in this chapter we consider the remaining two types of parameter required for cost effectiveness models: cost parameters and utility parameters.

### 7.1 Introduction

Having considered the range of distributions that are most frequently used for characterising effectiveness parameters and the uncertainty regarding their true values, in this chapter we consider the remaining two types of parameter required for cost effectiveness models: cost parameters and utility parameters. Section 7.2 considers distributions for cost parameters; specifically the LogNormal and the Gamma distributions. For each we describe how to fit the distribution and comment on how to choose which distribution to use. Section 7.3 considers distributions for utility parameters. Focusing on Normal, Beta and Gamma distributions, we consider how to fit these distributions and when to use each. We also consider some of the particular challenges that the properties of utility data pose for probabilistic analyses and some pragmatic solutions to these problems. Section 7.4 provides a summary of the key points covered in the chapter.

## 7.2 Distributions for Cost Parameters

Decision makers are often most interested in the costs of a new technology, how adopting the technology is likely to impact upon service use in other parts of the health and social care system and whether the technology might prove to be cost saving in the long run. As well as being the subject of great attention, there are often greater uncertainties affecting the cost estimates in an analysis than other parameters. There are substantial challenges in generalising the cost data from other health systems, reflecting the clinical practice and patient preferences in those systems, whilst using resource data from ‘domestic’ patients only inevitably increases the uncertainty due to smaller sample sizes. As a result, the appropriate characterisation of the uncertainty in cost parameters is important to provide decision makers with robust evidence that can withstand careful appraisal.

Cost data cannot be negative, and therefore the distribution we use for cost parameters must be strictly positive, ranging from zero to positive infinity. Hence the Beta distribution is not appropriate having a maximum value of one. The Normal distribution is unlikely to be appropriate due to the minimum value of minus infinity.

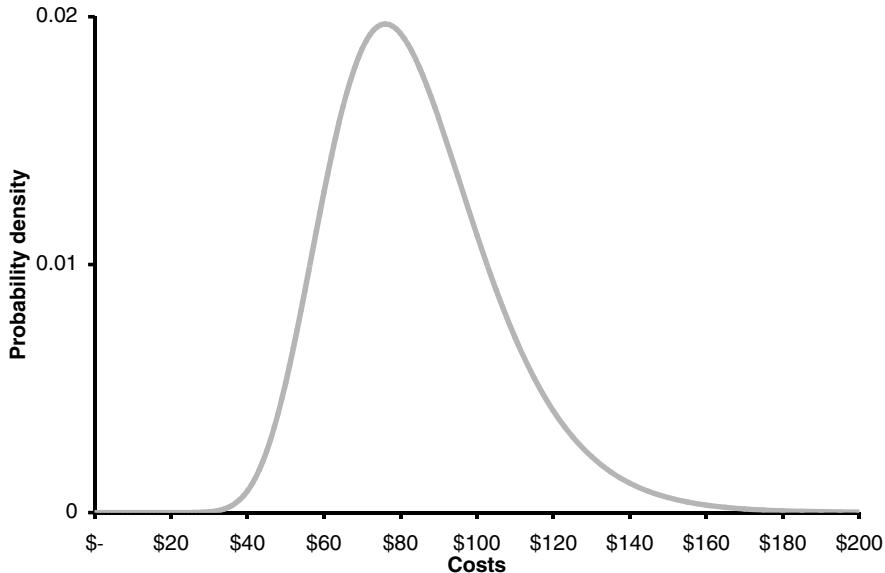
Empirically, we observe that health-care resource use tends to cluster around the mode value but a small number of observations take on very high values, many standard deviations away from the mode. For example, the majority of infants in neonatal intensive care units stay a matter of days; however, a very few infants will spend many months. Similarly, the majority of day surgery patients will be discharged home on the day of the operation, but very occasionally, perioperative complications lead to admission to the critical care unit and lengthy inpatient admissions. These rare observations have such large costs that they generate a meaningful difference between the mode, median and mean costs. This characteristic of resource use and cost data makes symmetrical distributions less appropriate.

The LogNormal and Gamma distributions are strictly positive and can deal with long-tailed data. For this reason these are very much the workhorse distributions for cost parameters in probabilistic analyses.

### 7.2.1 The LogNormal Distribution

The LogNormal distribution is formed by taking the exponential of a Normal distribution, i.e. the natural log of a Normal distribution. It is characterised by the parameters that characterise the Normal distribution,  $\mu$  and  $\sigma$  (mean and standard deviation). Figure 7.1 plots the LogNormal distribution with a mean of \$84 and a standard deviation of \$22.

If patient-level cost data is available, the first step is to take the natural logs of the patient-level data and fit a linear regression model to that data. The predicted mean and standard error from the regression model are the parameters required for the distribution in our cost effectiveness model. It is worth noting that when a single regression model is used to estimate the costs for more than one of the states in the model, there will be correlation between the costs for each health state. Such



**Fig. 7.1** A LogNormal distribution

correlation should be taken into account in the parameterisation of the model. We consider how to do this in Chap. 8 using the Cholesky Decomposition. For now, we will proceed as though the costs in each model health state are independent.

Normally, patient-level data are not available, and we must rely on summary statistics such as the mean, median and standard deviation. The following relationships are useful for calculating the parameters we need from such summary statistics.

If the log costs have mean  $\mu$  and standard deviation  $\sigma$ , then

$$\text{Median} = e^\mu$$

$$\text{Expected value} = E(x) = e^{\mu + \frac{1}{2}\sigma^2}$$

$$\text{Variance} = s^2 = \left( e^{\sigma^2} - 1 \right) e^{2\mu + \sigma^2}$$

### 7.2.1.1 LogNormal via Method of Moments I

If the published literature provides the expected value and standard deviation for the cost parameters, then we calculate  $\mu$  and  $\sigma$  for the LogNormal distributions as follows:

First calculate  $\sigma$  as

$$\sigma = \sqrt{\ln \left( 1 + \frac{s^2}{E(x)^2} \right)}$$

Having obtained  $\sigma$ , we can now calculate  $\mu$

$$\mu = \ln[\text{E}(x)] - \frac{1}{2}\sigma^2$$

### 7.2.1.2 LogNormal via Method of Moments II

If the published literature reports the expected value and the median, then we calculate  $\mu$  and  $\sigma$  as follows:

$$\mu = \ln(\text{Median})$$

and

$$\sigma = \sqrt{2(\ln[\text{E}(x)] - \mu)}$$

### 7.2.2 The Gamma Distribution

The Gamma distribution has two parameters,  $\alpha$  and  $\beta$ . The relationship between these two parameters and the mean and variance of the sample data is given by the following two functions:

$$\text{Expected value} = \text{E}(x) = \alpha\beta$$

and

$$\text{Variance} = s^2 = \alpha\beta^2$$

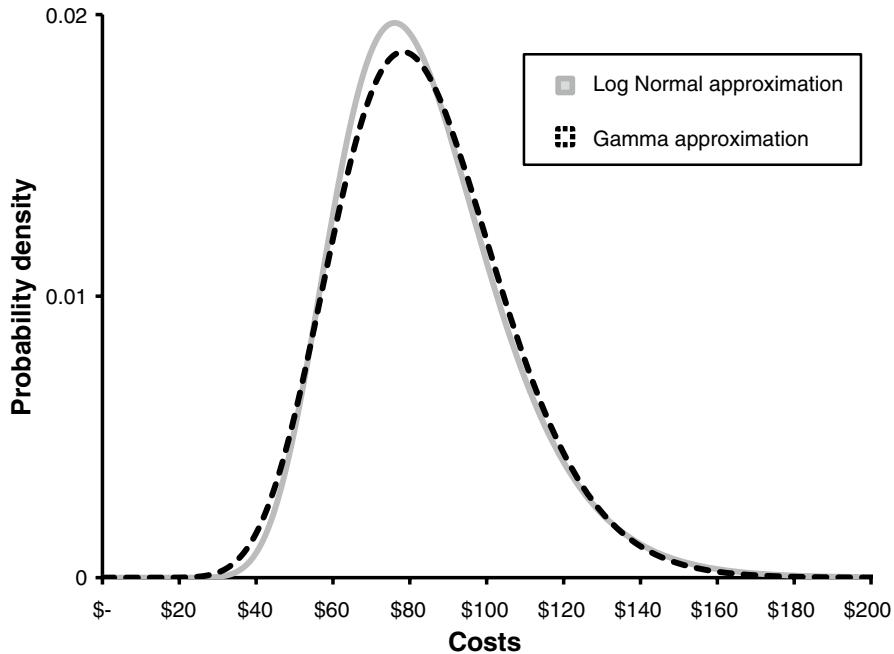
We use the method of moments to calculate  $\alpha$  and  $\beta$  from the expected value and variance of the sample data:

$$\beta = \frac{s^2}{\text{E}(x)}$$

and

$$\alpha = \frac{\text{E}(x)}{\beta}$$

The question that is likely already in your mind is: ‘How do I know when to use a LogNormal and when to use a Gamma distribution?’, and unfortunately there is not a simple answer. Fundamentally it comes down to which distribution fits your data



**Fig. 7.2** Comparison of LogNormal and gamma distributions

most closely. For data drawn from large samples, it will make little difference which distribution you choose. However, with smaller samples you will observe differences between the distributions. Figure 7.2 shows distribution plots for a LogNormal and Gamma distribution based upon the same expected value and standard deviation data. The range of the Gamma distribution is slightly larger, with a higher probability of observing values less than 50, whilst the LogNormal distribution is slightly taller, with a higher probability of observing values around the mode. To arrive at the decision of which distribution to use, ideally you will be able to compare the distribution to the actual data. When there is no clear rationale for choosing one over the other, then it would be sensible to run and analyse both to establish whether the results are sensitive to the choice of distribution. In Chap. 12 we will introduce you to Value of Information, which offers a quantitative approach to establishing whether the uncertainty associated with choosing one distribution over another has important implications for the decision the analyses attempt to inform.

### 7.3 Distributions for Utility Parameters

At the start of this section, it is useful to recap what utility data are and the role they play in cost effectiveness analysis. Utilities are the weights that are applied to life duration data to construct quality-adjusted life years (QALYs) (Williams et al. 1985).

QALYs are frequently used as the outcome measure for evaluation of health technologies and for reimbursement decisions. The outcome measure forms the denominator in the calculation of the increment cost effectiveness ratio (ICER) (Drummond et al. 1987), and thus the results of cost effectiveness analyses are often highly sensitive to small changes in the utilities. As a result, it is incumbent upon analysts to be extremely careful in the specification of utility parameters. We must provide clear justifications for our choices and, where possible, test and communicate the sensitivity of the model results to the alternative approaches.

Utilities are most frequently obtained from preference-based health-related quality of life (HrQoL) measures such as the EQ-5D, the SF-6D or the Health Utilities Index (Dolan 1997; Brazier et al. 2001; Feeny et al. 2002). These measures describe health states using domains and levels. A complete health state comprises one level descriptor from each domain (Brazier et al. 2007). For example, a health state in the EQ-5D instrument is constructed using one level descriptor each from mobility, self-care, usual activities, pain/discomfort and anxiety/depression (see Box 7.1). The EQ-5D-3L has three levels to each domain, whilst the newer EQ-5D-5L uses five levels.

It is worth noting that the three measures mentioned above are generic measures; that is, they are designed to measure the HrQoL of all people, from the very healthy to the very sick. Equivalent condition-specific HrQoL measures, with associated utility indices, have been developed for a range of conditions (Brazier and Tsuchiya

**Box 7.1: Example Health-Related Quality of Life Descriptive System:  
EQ-5D-3L Oemar and Oppe (2013)**

Mobility

- 1. I have no problems in walking about
- 2. I have some problems in walking about
- 3. I am confined to bed

Self-Care

- 1. I have no problems with self-care
- 2. I have some problems washing or dressing myself
- 3. I am unable to wash or dress myself

Usual Activities

- 1. I have no problems with performing my usual activities
- 2. I have some problems with performing my usual activities
- 3. I am unable to perform my usual activities

Pain/Discomfort

- 1. I have no pain or discomfort
- 2. I have moderate pain or discomfort
- 3. I have extreme pain or discomfort

Anxiety/Depression

- 1. I am not anxious or depressed
- 2. I am moderately anxious or depressed
- 3. I am extremely anxious or depressed

2010). Conceptually, the utility data from such instruments are on the same utility scale, and therefore the observations and recommendations in this chapter apply equally to utilities derived from such measures. That said, it is questionable whether it would be appropriate to combine utility data from condition-specific and generic measures in a single model. The legitimacy of this would require that the methods of the valuation studies that produced the utility data were the same, with specific consideration given to the anchors used in the valuation questionnaire, the valuation method and the characteristics of the respondents. Often decision makers will recommend a ‘reference case’ HrQol instrument in order to avoid problems of this sort. Further discussion of these issues is outside the scope of this text. Brazier et al. (2007) provides an excellent overview of the state of the art of the measurement and valuation of health.

Markov cost effectiveness models are constructed of health states, but these states will normally be defined in clinical terms (Briggs and Sculpher 1998), and it is unlikely that the descriptive system of HrQol measure will map perfectly onto the clinical health state descriptions. As a result the utility value for each health state in the model will likely be an aggregate of utilities for multiple HrQol states, reflecting the variation in patients’ experienced HrQol in each health state. The observed variation in the utility data for a given clinical health state may therefore capture both heterogeneity and uncertainty. The heterogeneity component can be thought of as the systematic variation that is due to patient subgroups within the population, which can be identified a priori. Heterogeneity should be addressed through subgroup analyses and not through the probabilistic sensitivity analysis. The utility parameter distribution should reflect the uncertainty around the expected value for the utility of the model health state, for a homogenous patient population. Ideally, as analysts we should reassure ourselves that the data we are using to characterise the parameter distribution does not contain distinct patient subgroups with distinct health utility values. In practice, the available data is often insufficient for this purpose, but it is still good practice to ask the question ‘Is there any reason to believe the observed variation in utilities contains a systematic component?’ Assuming we have data that is appropriate for describing the uncertainty in the expected value for the utility parameters, we then need to consider which distributions we should use to represent that uncertainty.

### 7.3.1 *Distributional Characteristics of the Utility Scale*

The utility scale in the QALY model has some very specific and unusual characteristics. First it is anchored and censored at 1.0. It is also anchored at zero but has a lower bound of minus infinity. To further complicate matters, for a number of widely used instruments, the lower bound has been artificially censored at -1. There are no distributions that match these characteristics (Dolan 1997; Brazier et al. 2001; Feeny et al. 2002). Therefore, whichever distribution we choose is going to be an approximation. In this context, four distributions are widely used to approximate the uncertainty in utility data: the Beta, Gamma, LogNormal and Normal distributions.

## 7.4 Characterising Uncertainty for Expected Utility Values Close to 1.0

The Beta distribution is intuitively attractive for the use with utility data because it also has a maximum value of 1.0. However, it is constrained to lie between 1.0 and 0.0, whereas many HrQol instruments have negative values (Dolan 1997; Feeny et al. 2002; McCabe et al. 2005). As a result the Beta distribution can be used most confidently when the expected value is close to 1.0 and the variance is small.

For example, if we had data showing the mean utility for a state was 0.85, with a standard deviation of 0.07, we could characterise the Beta distribution using the continuous specification of the Beta distribution:

$$\alpha + \beta = \frac{E(x)[1 - E(x)]}{s^2 - 1}$$

$$\alpha = E(x)[\alpha + \beta]$$

Inserting the values into the equations, we obtain

$$\alpha + \beta = \frac{[0.85 \times 0.15]}{0.0049 - 1} = 26.02$$

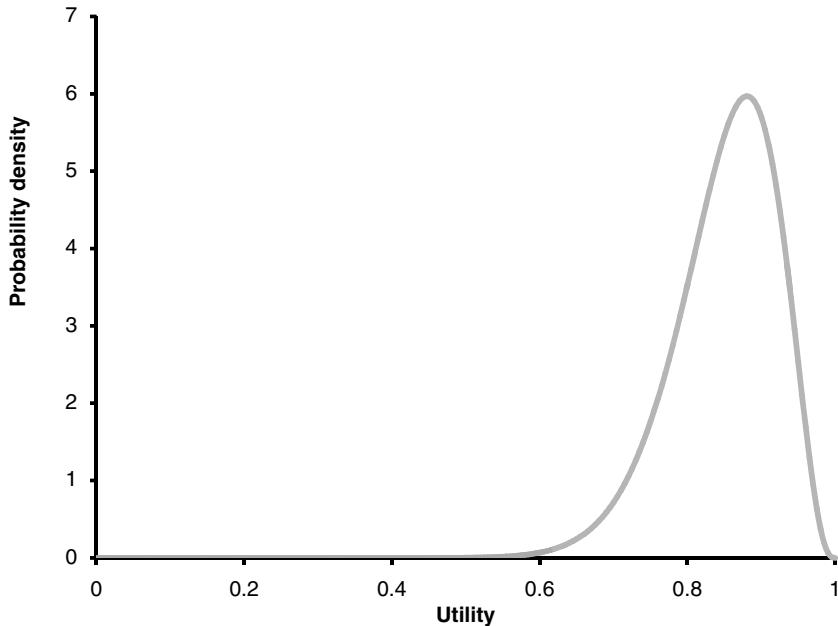
$$\alpha = 0.85 \times 26.02 = 22.12$$

$$\beta = [26.02 - 22.12] = 3.90$$

Figure 7.3 shows that there is little if any of the distribution below 0.5, and thus, the potential problem of censoring at zero is not a concern with this data. It is the combination of the proximity of the mean to 1.0 and the scale of the standard deviation that determines the suitability of the Beta distribution. Looking at a graph of the distribution is an excellent check that the Beta distribution is appropriate. If there are a substantial number of observations close to zero, it is credible that the use of the Beta distribution is censoring the underlying distribution.

### 7.4.1 Characterising Uncertainty for Expected Utility Values Away from 1.0

Unsurprisingly for many cost effectiveness analyses, there are multiple model health states that lie substantially away from 1.0. When we have reason to believe that the distribution needs to span zero, we need a distribution that can take negative values. How should we specify the uncertainty for these parameters? A widely used



**Fig. 7.3** Beta distribution;  $\alpha=22.12$ ,  $\beta=3.90$

technique is to transform the data from the utility scale to the disutility scale. The disutility scale is simply one minus utility, if

$$\begin{aligned} E(x_u) &= \mu, \sigma = s, \quad \text{where } x_u = \text{utility} \\ E(x_d) &= 1 - \mu, \sigma = s \quad \text{where } x_d = \text{disutility} \end{aligned}$$

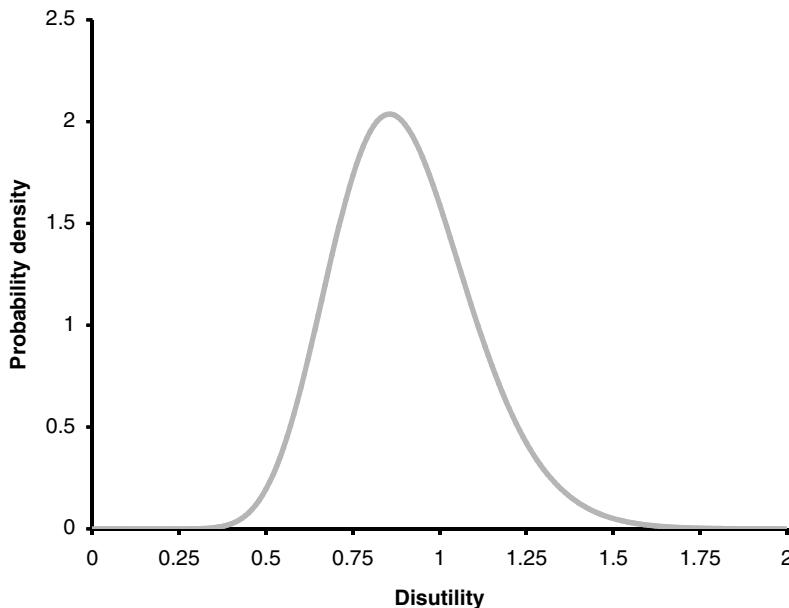
when  $x_u = 1$ ;  $x_d = 0$ .

Thus, the disutility scale is positive, and we can use distributions that are censored at zero but unbounded above zero. We have already used two such distributions in the consideration of the effectiveness and cost parameters: the LogNormal and the Gamma distribution.

Figure 7.4 shows the Gamma distribution for a utility distribution that spans zero on the disutility scale. The expected utility value is 0.10, with a standard deviation of 0.20.

$$\beta = \frac{s^2}{E(x)} = \frac{0.04}{0.9} = 0.0444$$

$$\alpha = \frac{E(x)}{\beta} = \frac{0.9}{0.0444} = 20.25$$



**Fig. 7.4** Gamma distribution;  $\alpha=20.25$ ,  $\beta=0.044$

The distribution is almost, but not quite, symmetrical and covers the range from 0.3 to 1.8 on the disutility scale.

Looking at the Gamma distribution below, it is reasonable to ask the question, ‘what is the advantage of a Gamma distribution over a Normal distribution?’. When we are dealing with health states that are substantially away from full health, using the Normal distribution is a perfectly reasonable strategy. Figure 7.5 is a Normal distribution with the same mean and standard deviation as Fig. 7.4, but on the disutility scale.

Whilst in this specific case the Normal distribution does not cause any problems, the fact that it is strictly unbounded makes it a less than satisfactory choice. In Fig. 7.6 we have plotted both the Gamma and the Normal distribution on the disutility scale. Plots like this can be used to assess whether one distribution is clearly a better fit to the observed data.

When we use the disutility scale, rather than the utility scale, for health utilities closer to one, the Normal distribution will not be appropriate. However, the LogNormal distribution may be appropriate as it is bounded at zero and unbounded above zero.

#### 7.4.2 *Logical Ordering for Utilities in Cost Effectiveness Models*

So far, we have considered health state utility parameters as though they are independent of other parameters in the model, including the utility parameters for other model health states. Frequently, even normally, there is a logical ordering to the

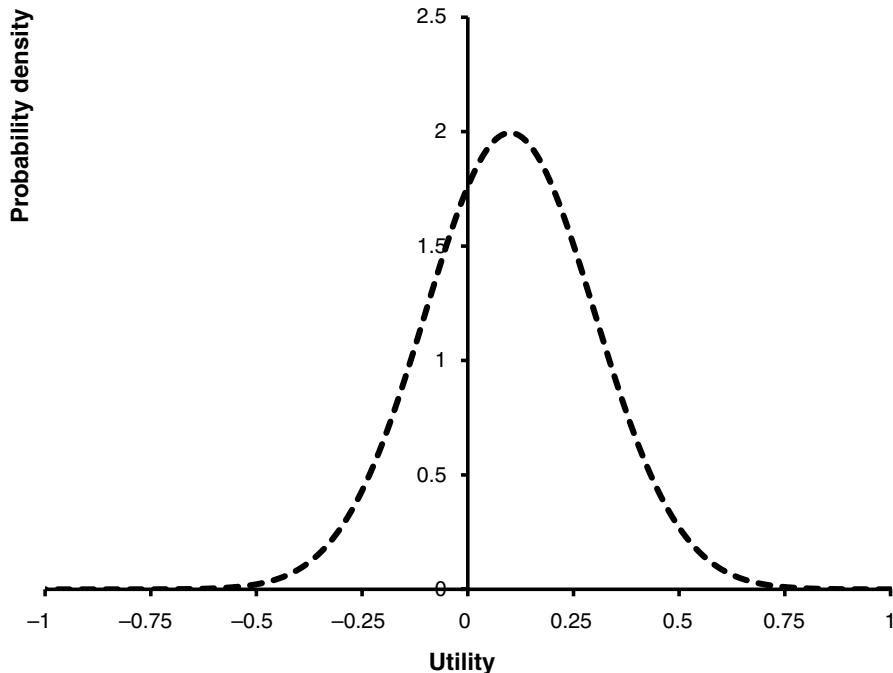


Fig. 7.5 Using a normal distribution for utilities far from 1.0

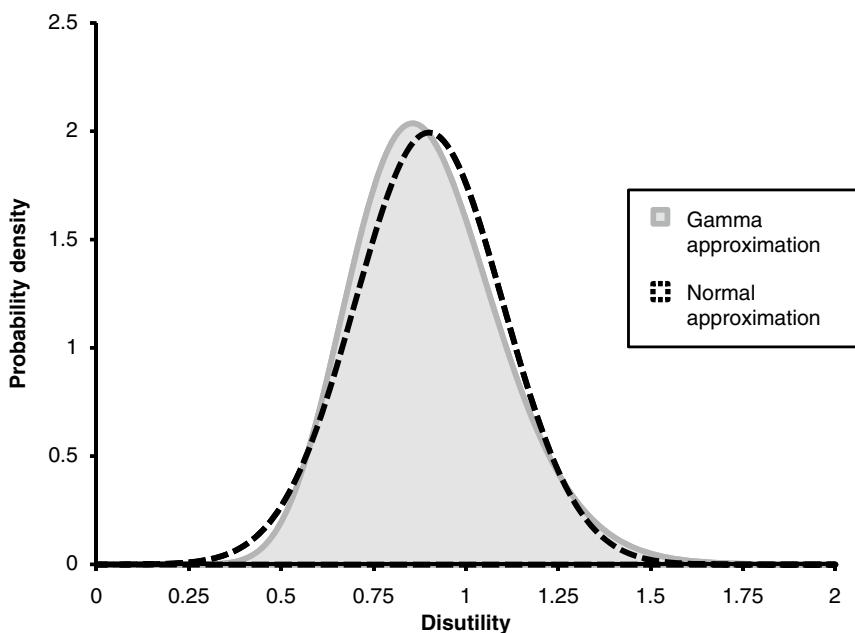


Fig. 7.6 Normal and gamma distribution approximations:  $x_d=0.9$ ,  $s_d=0.2$

utilities for each model health state. If each of the utility parameters is modelled independently, then some of the simulations may not adhere to the logical ordering and produce results that do not have face validity, e.g. that a cohort with untreated disease that progresses produces more QALYs than a cohort of patients who are treated and do not progress.

The simplest solution to this problem is to model utilities as decrements from the best health state. For example, if we had a simple three-state model – Stable disease, Progressive disease and Progressive disease with complications – we can be confident that Stable disease has a higher utility than Progressive disease and that Progressive disease has a higher utility than Progressive disease with complications. To preserve this logical ordering, we could characterise the utility for Stable disease using a Beta distribution, the utility decrement Progressive disease compared to Stable disease using a Gamma distribution and the utility decrement for Progressive disease with complications compared to Progressive disease using another Gamma.

If the health states in the model are all substantially away from one, allowing us to set aside the risk of sampling values greater than one, we could parameterise the utilities for all the health states as correlated Normal distributions using Cholesky Decomposition. We will explore this strategy in more detail in the next chapter.

#### ***7.4.3 Health State-Specific Side Effect Utility Decrements***

Often the technologies being evaluated are known to have side effects that impact upon HrQoL. However, the impact does not change the clinical health state; thus, there is a need to characterise an additional utility decrement. There are a number of issues to consider in parameterising the side effect utility decrements.

The greater the difference in HrQoL between the model states, the less plausible it is to assume that the decrement is the same for all states. At the same time, specification of uncertainty in the utility decrement needs to respect any logical ordering in order to avoid irrational results.

Health state decrements must be positive. The Normal distribution risks producing negative values and will likely not be appropriate for this reason. Similarly, for severe side effects the Beta distribution may not be appropriate as decrements greater than one may be feasible.

If there are multiple side effects, then consideration will need to be given as to the plausibility of assuming a shared utility decrement. When a shared decrement is not credible, then care will be required to clearly specify whether side effects have an additive, multiplication or some other interaction to specify the cumulative impact. As long as the probability of having each side effect is independent, then the uncertainty in the individual utility decrement can be parameterised with an independent distribution that provides inputs to an embedded utility function that is assumed to be certain in the same way that the clinical pathway is assumed to be certain.

## 7.5 Summary

- Cost data are strictly positive and tend to have a long tail – i.e. there is a small probability of extremely large mean costs.
- The choices for modelling the uncertainty in cost data tend to be confined to Gamma or LogNormal distributions.
- Whilst the Normal distribution may sometimes be a defensible choice, given the data required to parameterise a Normal distribution will allow the parameterisation of a LogNormal distribution, there is no strong argument to consider it.
- Both LogNormal and Gamma distributions can be parameterised using the methods of moments from expected value and standard deviation data that are usually reported in the published literature.
- Whenever possible the choice of distribution should be based upon consistency with the observed data.
- The results of cost effectiveness are frequently highly sensitive to changes in the utility parameters, making care in the specification of expected values and distributions particularly important.
- The nature of utility data, censored at one but unlimited below zero, makes it impossible to provide definitive guidance as to the ‘correct’ choice of distribution.
- Normal, Beta, LogNormal and Gamma distributions may all be appropriate in specific circumstances.
- Specifying parameters in disutility space can be an effective means of ensuring utilities never exceed 1.0.
- Care is required to ensure that the sampled utilities respect logical orderings implicit in the model structure. Specifying utility parameters as decrements from the utility of the best health state in the model is an effective means of doing this.
- Specifying distributions for utility decrements associated with adverse events in a model must also respect logical orderings within the model.

## References

- Brazier J, Tsuchiya A (2010) Preference-based condition-specific measures of health: what happens to cross programme comparability? *Health Econ* 9(2):125–129
- Brazier JE, Roberts J, Deverill M (2001) The estimation of a preference based measure of health from the SF-36. *J Health Econ* 21(2):271–292
- Brazier JE, Ratcliffe J, Salomon J, Tsuchiya A (2007) Measuring and valuing health benefits for economic evaluation. OUP, Oxford
- Briggs A, Sculpher M (1998) An introduction to Markov modeling for economic evaluation. *Pharmacoeconomics* 13(4):397–409
- Dolan P (1997) Modelling valuations for the Euroqol health states. *Med Care* 35(11):1095–1108

- Drummond MF, Stoddart GL, Torrance GW (1987) Methods for the economic evaluation of health care programmes, 1st edn. Oxford Medical Publications, Oxford
- Feeny D, Furlong W, Torrance GW, Zenglong Z, DePauw S et al (2002) Multi-attribute and single-attribute utility functions for the health utilities Index Mark 3 system. *Med Care* 40(2):113–128
- McCabe C, Stevens K, Roberts J, Brazier JE (2005) Health state values for the HUI2 descriptive system: results from a UK survey. *Health Econ* 14(3):418–431
- Oemar M, Oppe M (2013) EQ-5D-3L User Guide Euroqol Group. [http://www.euroqol.org/fileadmin/user\\_upload/Documenten/PDF/Folders\\_Flyers/EQ-5D-3L\\_UserGuide\\_2013\\_v5.0\\_October\\_2013.pdf](http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/EQ-5D-3L_UserGuide_2013_v5.0_October_2013.pdf)
- Williams A (1985) Economics of coronary artery bypass grafting. *BMJ* 291:736. <http://www.bmjjournals.com/content/291/6497/736.2>

# Chapter 8

## Correlated Parameters and the Cholesky Decomposition

**Abstract** So far, we have proceeded on the assumption that parameters are independent of each other. The effect of failing to take account of the correlation between parameters is to systematically overstate the nature and the degree of uncertainty in the outputs of a cost effectiveness model. In non-linear models, this could also bias the estimates of expected costs and outcomes. Hence, it is important to consider whether parameters are likely to be correlated and, so far as possible, take this into account. In this chapter, we introduce the most frequently used method for incorporating correlations into probabilistic analysis – Cholesky Decomposition.

### 8.1 Introduction

As we observed a number of times in the previous chapters, so far, we have proceeded on the assumption that parameters are independent of each other, i.e. there is no correlation between the values that any of the parameters in the model take. As shown in Fig. 6.9, the effect of failing to take account of the correlation between parameters is to systematically overstate the nature and the degree of uncertainty in the outputs of a cost effectiveness model. In non-linear models, this could also bias the estimates of expected costs and outcomes. Hence, it is important to consider whether parameters are likely to be correlated and, so far as possible, take this into account.

In this chapter, we introduce the most frequently used method for incorporating correlations into probabilistic analysis – Cholesky Decomposition. The Cholesky Decomposition deals with correlated parameters drawn from a multivariate Normal distribution. Prior to the detailed exposition, it is useful to consider what we mean by correlated parameters in a little more detail. This chapter is unavoidably more mathematically demanding than the other chapters, but it will reward the effort required to work through it. By the end of the chapter, you will have an understanding of not only the theoretical material but also the tractable stepwise mechanism for implementing the Cholesky Decomposition in practice. Section 8.2 explains the concept of correlated parameters and why it is important to take account of the correlation in probabilistic analyses. Section 8.3 provides a formal characterisation of the set of correlated parameters, which Sect. 8.4 builds upon to describe the Cholesky Decomposition. Section 8.5 then discusses how to extend the decomposition for

larger numbers of parameters, and Sect. 8.6 briefly considers the interpretation of the information contained in the Cholesky Decomposition. The final section provides a summary of the key points in this chapter.

## 8.2 Correlated Parameters

In most cases, it is safe to assume that parameters are uncorrelated. That is, the value that one parameter takes will give us no information about the second parameter. For example, we might expect that the transition probability of moving from one state to another would be unrelated to the cost of spending a period in either of the states concerned. In other cases though, there may be sound reasons to expect that multiple parameters will be related, and here, it is important that this correlation is reflected in the model. This might occur either because we expect that two parameters will be related or because multiple parameters are used to characterise an underlying process.

A common case where parameters are related is where there is an element of severity within the model. Here, it is often the case that more severe ill health will be related to both higher costs and lower health-related quality of life. If we take the latter example, suppose we have utilities attached both to a severe disease state and to a more moderate disease state; we might expect lower figures for the first than in the second. If health in a moderate disease state is much lower than expected, then failing to incorporate this relationship could mean that model considers some cases that we don't expect to be credible. In this case, it is possible to also characterise this relationship by modelling the utility in one state and a second figure providing the difference between the states. It *may* also be possible to treat these 'baseline' and 'difference' figures as being independent in some cases.

It is more difficult, however, to avoid correlated parameters where multiple parameters are used to characterise an underlying process – for example, where this relates to costs or survival figures that change over time. Suppose we were interested in the costs that are incurred within a particular health state, where these costs are affected by how long one spends in the hospital. It is often the case that in these cases, we would expect some fixed upfront costs to be incurred followed by an additional amount that depends on the length of stay. If we had some data that could inform our estimates, we could analyse this by regressing the number of days in the hospital on the natural logarithm of costs. In this case, the constant term provides an estimate of the (log-)fixed costs, and the coefficient on length of stay tells us how much (log-)costs vary for each extra day in the hospital. These estimates both have confidence intervals that indicate the variability in the parameters, but these are not independent.

For the sake of argument, suppose that average costs in our dataset were about \$1,500, of which the regression estimates that the fixed costs were about \$1,400 and the remainder is related to length of stay. In this case, the difference between these figures suggests that length of stay is probably not that important when estimating

costs, since the vast majority of the costs are incurred up front. However, this estimate might be wrong and suppose instead that the fixed cost component was \$700. If we only changed the fixed cost component (which falls by \$700), the average estimated costs here would fall from around \$1,500 in the first case to around \$800. Clearly, changing only the fixed cost component would mean that we no longer accurately model our original data. Instead, it is necessary that as the estimated fixed cost component goes down, the model also has to allow (log-)costs to rise more quickly for each additional day in hospital. This allows the model to keep the average estimated cost from the model equal to the average cost that was actually observed. The two parameters are negatively correlated. In a similar way, if we were talking about a survival curve, a higher failure over the initial period of a model would need to be balanced by assuming a lower failure in subsequent periods in order to retain the expected failure rate within an observed dataset.

More broadly, when parameters are correlated, this means that the information that determines the value of one parameter will also partially determine the value of the other. For positively correlated pairs of parameters, we will tend to observe both parameters moving in the same direction. For negatively correlated pairs of parameters, one parameter will tend to take a higher than average value whilst the other will take a lower than average value. Correlation, in this setting, relates to how deviations from an average value are distributed – i.e. whether the deviations tend to move in the same or opposite directions.

When we draw parameters from distributions for correlated parameters; once we draw the value for one parameter, for subsequently drawn parameters, it is as if we know the first parameter value for certain. At this point, the uncertainty in the other parameters will be reduced, since we are now interested in the distribution for these parameters *conditional* on the parameter we already know. With a correlated set of parameters, each additional parameter that is already drawn increases the knowledge we have about the parameters yet to be drawn.

### 8.3 Defining a Set of Correlated Parameters

More formally, when we sample correlated parameters from distributions, we assume that they are drawn from a multivariate Normal distribution (also called a multivariate Gaussian distribution). We need to assume here that we have variability in all of these variables and none of them can be represented as a simple linear combination of any of the others. (This is hardly unreasonable, as there is no point drawing from a distribution with no uncertainty, and if we had a linear combination, we would not need to draw a parameter value for it since we could just calculate it based on other variables.)

Whereas a univariate (one-variable) Normal distribution is distributed with a mean and standard deviation (or variance), a multivariate Normal distribution is distributed with a vector of means (one for each variable) and a variance-covariance matrix. This variance-covariance matrix contains information on both the total

uncertainty in each variable and how the uncertainties between each item are related. For a three-variable case where we need to draw from  $x_1, x_2$  and  $x_3$ , the vector of means and variance-variance matrix will be as follows:

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix}, \quad \boldsymbol{\Omega} = \begin{bmatrix} \sigma_1^2 & \rho_{1,2}\sigma_1\sigma_2 & \rho_{1,3}\sigma_1\sigma_3 \\ \rho_{1,2}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{2,3}\sigma_2\sigma_3 \\ \rho_{1,3}\sigma_1\sigma_3 & \rho_{2,3}\sigma_2\sigma_3 & \sigma_3^2 \end{bmatrix}$$

where  $\sigma_i$  is the standard deviation of  $x_i$  and  $\rho_{i,j}$  is the covariance between  $x_i$  and  $x_j$ .

When we use a multivariate Normal distribution, each of the individual variables is itself drawn from a Normal distribution, and if we have no additional information about the distribution, then this is the only thing we can do. So, if  $x_1$  is the first variable on which we will draw information, then its distribution will be

$$x_1 \sim \text{Normal}(\mu_1, \sigma_1^2)$$

Now in order to make things much easier for ourselves, we will normally want to concentrate on how much each variable differs from its mean value. That is, how many standard deviations from the mean are we talking about? Usefully, there is a standard type of variable that measures outcomes in this way. A standard normal variate is a random variable that is drawn from a Normal distribution with mean 0 and standard deviation 1. If we characterise  $x_1-x_3$  using normal variates  $y_1-y_3$ , then we can write all three variables as

$$\begin{aligned} x_1 &= \mu_1 + \sigma_1 y_1, \quad x_1 \sim \text{Normal}(\mu_1, \sigma_1^2) \\ x_2 &= \mu_2 + \sigma_2 y_2, \quad x_2 \sim \text{Normal}(\mu_2, \sigma_2^2) \\ x_3 &= \mu_3 + \sigma_3 y_3, \quad x_3 \sim \text{Normal}(\mu_3, \sigma_3^2) \end{aligned}$$

Now if the  $x$  variables are correlated, then so too will be the  $y$  variates – and the variance-covariance matrix given above reflects this system of equations. However, it is more complicated to draw correlated variables from a distribution directly, and it would be easier if we could reformulate this system of equations in terms of three uncorrelated variates,  $z_1-z_3$ .

For our first variable  $x_1$ , we can write this very easily as

$$x_1 = \mu_1 + \sigma_1 z_1$$

Looking at the equation above, it is pretty clear that we have to set  $y_1 = z_1$ . That is, to find our first parameter, we just use the first of our uncorrelated normal variates – and let this take the value  $\hat{z}_1$ . Once we know this value, we can write

$$x_1 \mid \hat{z}_1 = \mu_1 + \sigma_1 \hat{z}_1$$

In terms of notation, a conditional distribution is reflected by a vertical line | that separates the thing we are interested in (in the equation ‘ $x_1$ ’) and the thing that we know about (in the equation ‘ $\hat{z}_1$ ’).

For our second variable, we know that the unconditional distribution (where we don’t know anything about the value of the variables) has a mean value of  $\mu_2$  and a variance of  $\sigma_2^2$ . However, the conditional distribution for  $x_2$  will depend on the value we’ve drawn for  $x_1$  and in particular on how far away it is from its mean value (i.e. on  $\hat{z}_1$ ). Our expected value for  $x_2$  will be

$$E(x_2 | \hat{z}_1) = \mu_2 + a\hat{z}_1.$$

Now before we knew something about  $z_1$ , the variance of  $x_2$  was  $\sigma_2^2$ . Once we know  $\hat{z}_1$ , some of the uncertainty in  $x_2$  is resolved. We use the second uncorrelated normal variate  $z_2$  to reflect the rest of this uncertainty. So here,  $x_2$  is found by drawing value from  $z_2$  so that the distribution for  $x_2$  becomes

$$x_2 | \hat{z}_1 = \mu_2 + a\hat{z}_1 + bz_2$$

At this stage, do not worry about what  $a$  and  $b$  are, since we will define these later. This conditional distribution will have a different variance than the unconditional one, since only the last term on the right-hand side is unknown. Then this means that there will now be *less* uncertainty left in the conditional distribution.

$$\text{Var}(x_2 | \hat{z}_1) = b^2$$

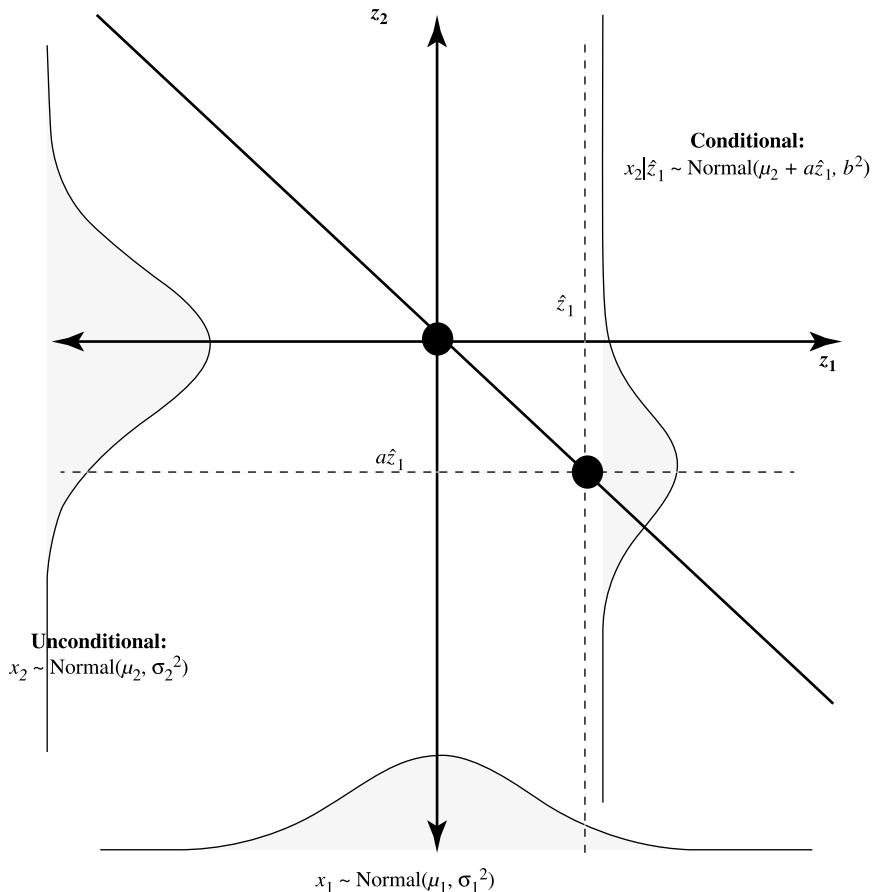
Figure 8.1 provides an illustration of how these two distributions are related in terms of the random variates. Here, the two random normal variates ( $x_1$ ,  $x_2$ ) appear on the axes with the axes intersecting at the unconditional means. The unconditional distribution for  $x_1$  appears on the bottom of the vertical axis, and the unconditional distribution for  $x_2$  appears at the left of the horizontal axis. Once we sample  $\hat{z}_1$ , we know the value for  $x_1$  and can predict a conditional distribution for  $x_2$ . In the illustrated example, the two distributions pictured are negatively correlated, so that when  $\hat{z}_1$  is positive, we have selected a value for  $x_1$  that is higher than its mean value and now expect a value for  $x_2$  that is lower than its mean value. The conditional distribution pictured for  $x_2 | \hat{z}_1$  also displays the *smaller* variance we observed because some of the uncertainty is resolved.

In practice, the value we obtain for  $x_2$  will depend upon the value sampled for  $z_2$ ; we can call this  $\hat{z}_2$ . Now here, we can write

$$x_2 | \hat{z}_1, \hat{z}_2 = \mu_2 + a\hat{z}_1 + b\hat{z}_2$$

In a similar way, our third parameter  $x_3$  will have a conditional mean that is a function of the values found for the previous two normal variates ( $\hat{z}_1, \hat{z}_2$ ):

$$E(x_3 | \hat{z}_1, \hat{z}_2) = \mu_3 + c\hat{z}_1 + d\hat{z}_2$$



**Fig. 8.1** Correlated random variates

If this is drawn using another standard normal variate,  $z_3$ , then

$$x_3 = \mu_3 + c\hat{z}_1 + d\hat{z}_2 + e z_3$$

So that

$$\text{Var}(x_3 | \hat{z}_1, \hat{z}_2) = e^2$$

To sample a value for  $x_3$ , we draw from our third uncorrelated normal variate to find  $\hat{z}_3$ . Here,

$$x_3 | \hat{z}_1, \hat{z}_2, \hat{z}_3 = \mu + c\hat{z}_1 + d\hat{z}_2 + e\hat{z}_3$$

For all three of the variables we have considered here, we have followed a similar process. Here, we start by considering the impact that any previously drawn variates have on the expected value of a variable, before drawing the next variate to deal with the current variable of interest. In practice, this means that the first variable drawn relies on the first variate only, the second variable relies on the first two variates and so on. When doing this, we have used five values  $a - e$  when defining our variables with uncorrelated normal variates but have not suggested any values for them. The way that we have approached this problem allows us to use a mathematical technique known as the Cholesky Decomposition.

## 8.4 The Cholesky Decomposition

We have now defined all three of our variables  $x_1 - x_3$  using draws taken from three uncorrelated normal variates  $z_1 - z_3$ . (As the equations hold irrespective of which values we draw, we will refer to these variates from now on as  $z_1 - z_3$  rather than as  $\hat{z}_1 - \hat{z}_3$ .) This means that we are now again talking about the same unconditional distributions as we began with. If we write this in matrix terms, our variables become

$$\mathbf{x} - \boldsymbol{\mu} = \mathbf{L}\mathbf{z}$$

where

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix}, \quad \mathbf{L} = \begin{bmatrix} \sigma_1 & 0 & 0 \\ a & b & 0 \\ c & d & e \end{bmatrix} \text{ and } \mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$$

Written this way, the critical part of this system of equations is the  $\mathbf{L}$  matrix, which is a lower triangular matrix with real and positive diagonal entries (i.e. the top left, middle centre, bottom right)<sup>1</sup>.

Now the variance-covariance matrix for system  $\mathbf{x} - \boldsymbol{\mu}$  will equal

$$\boldsymbol{\Omega} = \mathbf{L}\mathbf{L}^T \begin{bmatrix} \sigma_1 & 0 & 0 \\ a & b & 0 \\ c & d & e \end{bmatrix} \begin{bmatrix} \sigma_1 & a & c \\ 0 & b & d \\ 0 & 0 & e \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & a\sigma_1 & c\sigma_1 \\ a\sigma_1 & a^2 + b^2 & ac + bd \\ c\sigma_1 & ac + bd & c^2 + d^2 + e^2 \end{bmatrix}$$

---

<sup>1</sup> Within the equations defined to date, the off-diagonal values ( $a, c$  and  $d$ ) could take positive, negative, or zero values. (For an uncorrelated set of variables, all these off-diagonal values would equal zero.) In the matrix  $b$  and  $e$  cannot be equal to zero because that would mean either that there was no uncertainty (if we had an uncorrelated set of variables), or that we could represent at least one variable as a simple linear combination of the others (if we had a correlated set of variables). If we found this to be the case, we would stop and re-run the Cholesky Decomposition without this variable, and then calculate its value afterwards.

This  $\mathbf{L}$  matrix is known as the Cholesky Decomposition of the variance-covariance matrix  $\Omega$ . The variance-covariance matrix here will be the same as the one defined earlier for the correlated system of variables.

$$\Omega = \begin{bmatrix} \sigma_1^2 & \rho_{1,2}\sigma_1\sigma_2 & \rho_{1,3}\sigma_1\sigma_3 \\ \rho_{1,2}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{2,3}\sigma_2\sigma_3 \\ \rho_{1,3}\sigma_1\sigma_3 & \rho_{2,3}\sigma_2\sigma_3 & \sigma_3^2 \end{bmatrix}$$

As we have uncertainty in all of our variables<sup>2</sup>, each variance-covariance matrix can be represented by only one  $\mathbf{L}$  matrix of this type for each (ordered) set of variables. So, if the computer package that you are using to obtain regression results for parameters can provide you with a Cholesky Decomposition, then this matrix can be displayed and the values in this matrix can be used to define your variables. The process to do this is now outlined.

So, we now have two definitions for the variance-covariance matrix that should be identical if they are to represent the same set of correlated variables. This means that every single corresponding cell of the matrices should contain the same value:

$$\begin{bmatrix} \sigma_1^2 & a\sigma_1 & c\sigma_1 \\ a\sigma_1 & a^2 + b^2 & ac + bd \\ c\sigma_1 & ac + bd & c^2 + d^2 + e^2 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \rho_{1,2}\sigma_1\sigma_2 & \rho_{1,3}\sigma_1\sigma_3 \\ \rho_{1,2}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{2,3}\sigma_2\sigma_3 \\ \rho_{1,3}\sigma_1\sigma_3 & \rho_{2,3}\sigma_2\sigma_3 & \sigma_3^2 \end{bmatrix}$$

To make these matrices identical, we need to start defining the terms that we do not know on the left-hand side (i.e.  $a, b, c, d, e$ ) using the items that we can observe within the variance-covariance matrix on the right. To do this, we will go row by row.

On the first row, the top left-hand cell in the variance-covariance matrix gives the variance for  $x_1$ , and this is identical in both cases; we don't need to do anything here. The covariance between  $x_1$  and  $x_2$  (top centre cell) is  $a\sigma_1$  on the left-hand side and  $\rho_{1,2}\sigma_1\sigma_2$  on the right-hand side; both definitions will be identical so long as

$$a = \rho_{1,2}\sigma_2$$

If we look at the covariance between  $x_1$  and  $x_3$  (top right cell), we can also define  $c$ . For the left-hand side, this is equal to  $c\sigma_1$  and  $\rho_{1,3}\sigma_1\sigma_3$  on the right. Again, these definitions will be identical so long as

$$c = \rho_{1,3}\sigma_3$$

---

<sup>2</sup> So that the variance-covariance matrix is positive definite; if one or more variables had no uncertainty or a variable could be represented as a linear function of the others, it would have been positive semi-definite. However, both of these cases have been excluded.

Now, since the variance-covariance matrix is symmetrical, then we solved the left-hand column at the same time as we solved the top row. When we look at the middle centre cell, we consider the variance of  $x_2$ . On the left-hand side, this is equal to  $a^2 + b^2$  and on the right-hand side  $\sigma_2^2$ . In order to find a value for  $b$ , we can take advantage of the fact that we already have a value for  $a$ . Here,

$$a^2 + b^2 = \sigma_2^2$$

$$b = \sqrt{\sigma_2^2 - a^2}$$

Likewise, the middle right cell provides the covariance between  $x_2$  and  $x_3$ . The cell on the left-hand side contains three things that we already have values for ( $a, b, c$ ) and one unknown ( $d$ ), whilst the right-hand side contains  $\rho_{2,3}\sigma_2\sigma_3$ . We can solve for the unknown quantity  $d$  as

$$ac + bd = \rho_{2,3}\sigma_2\sigma_3$$

$$d = \frac{\rho_{2,3}\sigma_2\sigma_3 - ac}{b}$$

As the left and middle cells of the bottom row are the same as the top two cells in the right-hand row, this only leaves one cell still to solve. The bottom right cell contains the variance of variable  $x_3$ , and again we take advantage of the fact that we already have values for all but one of the items. In this case, we need to finally solve

$$c^2 + d^2 + e^2 = \sigma_3^2$$

$$e = \sqrt{\sigma_3^2 - c^2 - d^2}$$

So now, we can derive values for all five of the unknown quantities  $a - e$ . So long as the values for these variables are derived in the same order as presented here, we have provided a solution to this case (with three variables).

Once you have figures for the Cholesky Decomposition, it is a very quick process to introduce correlated parameters into a model. To do this, start by defining the number of uncorrelated normal variates that you require. From here, the correlated parameters will equal the mean value for the parameter plus a linear combination of these variates, where the weights placed on each variate are given by the figures  $a - e$  above.

## 8.5 What If I Need a Cholesky Decomposition for a Different Number of Variables?

If you need a Cholesky Decomposition for two variables, then the solution is somewhat simpler than provided in our example here. If we required a system of equations with only two variables, then our variance-covariance matrix would consist of only the two leftmost/topmost elements of the one we have dealt with. In this case, we would have a system of equations:

$$\mathbf{x} - \boldsymbol{\mu} = \mathbf{L}$$

$$\text{where } \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \mathbf{L} = \begin{bmatrix} \sigma_1 & 0 \\ a & b \end{bmatrix} \text{ and } = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

Here, the solutions for the Cholesky Decomposition are the same as those presented above. That is, we would again have

$$a = \rho_{1,2}\sigma_2$$

$$b = \sqrt{\sigma_2^2 - a^2}$$

If we instead required a Cholesky Decomposition for a larger number of variables, then the process is again very similar but expands upon the existing example. Appendix 8.1. describes the process for applying the Cholesky Decomposition for up to seven correlated parameters.

## 8.6 Interpreting the Cholesky Decomposition

As a form of matrix algebra, it is probably fair to say that the Cholesky Decomposition does not lend itself to very many ‘natural’ interpretations. For a given ordering of the variables that we are interested in, the Cholesky Decomposition breaks down the degree to which each successive variable is related to all proceeding ones. Properly interpreted, the decomposition can give some potentially useful information about the degree to which variables are related.

Within the decomposition, these interpretations are based on both the sign and the magnitude of some of the entries within each row. It is important to only try to interpret the decomposition within each row, as these rows relate to the uncertainty within one specific variable.

The size of the diagonal entries versus the other entries in a row tells us something about the degree to which the variability in a correlated variable is unique

within all those considered to date. To see this, let's consider the last variable of the three-variable example (the final row of the Cholesky Decomposition). When we look at the variance-covariance matrix, the variance of this final variable is given by

$$c^2 + d^2 + e^2 = \sigma_3^2$$

Here, the bigger is  $e$  relative to  $c$  and  $d$ , the less this third variable is related to the others. Suppose the values of  $c$ ,  $d$  and  $e$  are 0.474, -1.061 and 0.948, so that the variance of this variable is 2.249 (i.e. as  $c^2 + d^2 + e^2$ ). Around 10 % of the variation in the third variable (i.e.  $(0.474)^2/2.249$ ) comes from the same source as the variation of the first variable, with around 50 % from additional variation in the second variable (i.e.  $(-1.061)^2/2.249$ ). This suggests that only about 40 % of the variability in the third variable would be left to discover if we knew the value in the first two variables. This sort of result can be potentially useful when considering the value of future research, since this may give an indication of the degree to which resolving variability elsewhere might affect our confidence in predicting the value of this variable.

The sign of the coefficients may also be informative. A negative sign on the off-diagonal coefficients suggests that two sources of variability are negatively correlated, with a positive sign suggesting a positive correlation. This allows a slightly different view of correlation than provided by the variance-covariance matrix, since the variance-covariance matrix talks instead about the uncertain in different variables, rather than the sources of variability unique to each variable in turn.

## 8.7 Summary

- Probabilistic analyses frequently assume that model parameters are independent, i.e. that the value each parameter takes is independent of the value taken by any other parameter.
- Under certain circumstances, the assumption of independence is not defensible; for example, the parameters of survival functions will be correlated.
- The Cholesky Decomposition allows us to characterise the correlation between parameters by disaggregating the predicted value into an independent component and a component determined by the value taken by one or more different parameters.
- The decomposition matrix rows provide insight into the magnitude and direction of the interaction between each parameter.
- Failure to take account of the correlation between parameters will lead to an understatement of the uncertainty in the outputs from the model.

## Appendix

### Appendix 8.1: Extending the Cholesky Decomposition for More Than Three Correlated Parameters

There is no limit to the number of correlated variables we might want to model. However, most of the time it is pretty rare to see more than four or five variables correlated together in models. Since this is normally going to be more than enough, this appendix looks at the case of seven correlated variables.

Seven correlated variables ( $x_1 - x_7$ ) can be modelled as a function of the mean values and seven random variates  $z_1 - z_7$  multiplied by a matrix containing the values  $a$  to  $\mathbf{ae}$ .

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \\ \mu_5 \\ \mu_6 \\ \mu_7 \end{bmatrix} + \begin{bmatrix} \sigma_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ a & b & 0 & 0 & 0 & 0 & 0 \\ c & d & e & 0 & 0 & 0 & 0 \\ f & g & h & j & 0 & 0 & 0 \\ k & m & n & p & q & 0 & 0 \\ r & s & t & u & v & \mathbf{x} & 0 \\ y & z & \mathbf{aa} & \mathbf{ab} & \mathbf{ac} & \mathbf{ad} & \mathbf{ae} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \\ z_7 \end{bmatrix}$$

If you need six variables ( $x_1 - x_6$ ), then there would only be six random variates ( $z_1 - z_6$ ) and the matrix would only need the top six rows and leftmost six columns and we would need to find values  $a$  to  $x$ . With five variables, we would again have the ‘top-leftmost’ five rows and five columns and need to find values  $a$  to  $q$ , and so on.

Values for  $a$  to  $\mathbf{ae}$  can be found by solving the following equations in sequence. Note that each parameter uses either items from the variance-covariance matrix above and other quantities that we have already solved.

The only difference in notation between the Chap. 8 and this appendix is that we have removed the commas when referring to correlations, as we wish to keep the later formulas as short as we can. For example, the correlation between the first and second variables would have been  $\rho_{1,2}$  in the Chapter but appears as  $\rho_{12}$  here.

Two variables:

$$a = \rho_{12}\sigma_2$$

$$b = \sqrt{\sigma_2^2 - a^2}$$

Three variables – use  $a$  to  $b$  above and

$$c = \rho_{13}\sigma_3$$

$$d = \frac{\rho_{23}\sigma_2\sigma_3 - ac}{b}$$

$$e = \sqrt{\sigma_3^2 - c^2 - d^2}$$

Four variables – use  $a$  to  $e$  above and

$$f = \rho_{14}\sigma_4$$

$$g = \frac{\rho_{24}\sigma_2\sigma_4 - af}{b}$$

$$h = \frac{\rho_{34}\sigma_3\sigma_4 - cf - dg}{e}$$

$$j = \sqrt{\sigma_4^2 - f^2 - g^2 - h^2}$$

Five variables – use  $a$  to  $j$  above and

$$k = \rho_{15}\sigma_5$$

$$m = \frac{\rho_{25}\sigma_2\sigma_5 - ak}{b}$$

$$n = \frac{\rho_{35}\sigma_3\sigma_5 - ck - dm}{e}$$

$$p = \frac{\rho_{45}\sigma_4\sigma_5 - fk - gm - hn}{j}$$

$$q = \sqrt{\sigma_5^2 - k^2 - m^2 - n^2 - p^2}$$

Six variables – use  $a$  to  $q$  above and

$$r = \rho_{16}\sigma_6$$

$$s = \frac{\rho_{26}\sigma_2\sigma_6 - ar}{b}$$

$$t = \frac{\rho_{36}\sigma_3\sigma_6 - cr - ds}{e}$$

$$u = \frac{\rho_{46}\sigma_4\sigma_6 - fr - gs - ht}{j}$$

$$v = \frac{\rho_{56}\sigma_5\sigma_6 - kr - ms - nt - pu}{q}$$

$$x = \sqrt{\sigma_6^2 - r^2 - s^2 - t^2 - u^2 - v^2}$$

Seven variables – use  $a$  to  $x$  above and

$$y = \rho_{17}\sigma$$

$$z = \frac{\rho_{27}\sigma_2\sigma_7 - ay}{b}$$

$$\mathbf{aa} = \frac{\rho_{37}\sigma_3\sigma_7 - cy - dz}{e}$$

$$\mathbf{ab} = \frac{\rho_{47}\sigma_4\sigma_7 - fy - gz - h \bullet \mathbf{aa}}{j}$$

$$\mathbf{ac} = \frac{\rho_{57}\sigma_5\sigma_7 - ky - mz - n \bullet \mathbf{aa} - p \bullet \mathbf{ab}}{q}$$

$$\mathbf{ad} = \frac{\rho_{67}\sigma_6\sigma_7 - ry - sz - t \bullet \mathbf{aa} - u \bullet \mathbf{ab} - v \bullet \mathbf{ac}}{x}$$

$$\mathbf{ae} = \sqrt{\sigma_7^2 - y^2 - z^2 - \mathbf{aa}^2 - \mathbf{ab}^2 - \mathbf{ac}^2 - \mathbf{ad}^2}$$

# **Chapter 9**

## **Building a Markov Cost Effectiveness Model in Excel**

**Abstract** Within previous chapters, we have introduced you to Markov models. The objective of this and the next chapter is to provide you with an opportunity to bring together much of the material we have covered to this point by building a Markov cost effectiveness model. Working your way through this chapter, you will construct the deterministic model. We illustrate the model building process using Microsoft Excel. The model chosen is a stylised representation of a chronic disease which is characterised by periods of controlled disease with periodic disease flairs and an accumulation of long-term disability. This type of structure would suit conditions such as rheumatoid arthritis, cardiovascular disease and diabetes. We are going to consider the simplest form of cost effectiveness analysis – a comparison of two alternative treatments in terms of their costs and outcomes.

### **9.1 Introduction**

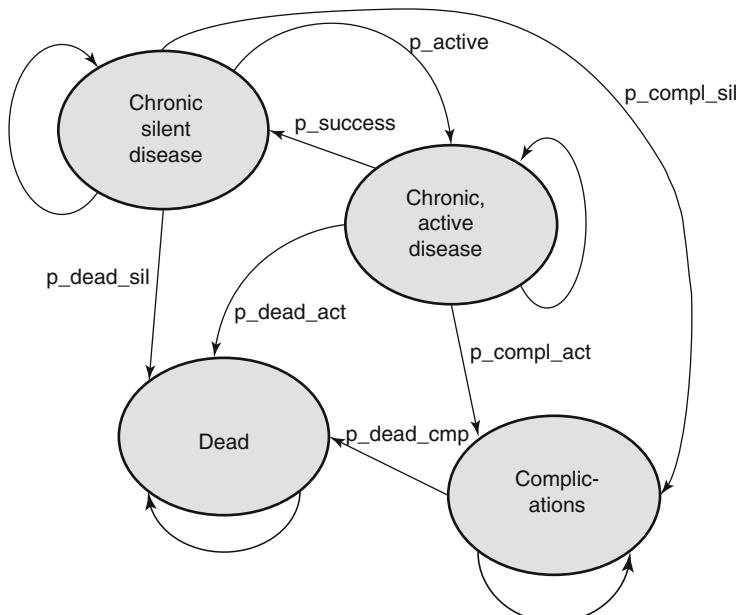
The objective of this and the next chapter is to provide you with an opportunity to bring together much of the material we have covered to this point by building a cost effectiveness model. Working your way through this chapter, you will construct the deterministic model. Chapters 10, 11 and 12 will then provide an opportunity to modify the model to provide probabilistic results and then produce a range of outputs that reimbursement decision makers may find helpful. Section 9.2 describes the conceptual model that you are going to build. Section 9.3 provides some useful tips for building models in Excel and requires you to build the transition. Section 9.4 deals with constructing the parameter table and the fields for capturing the Markov trace data. Section 9.5 then requires you to programme your model to generate the Markov trace data. In Sect. 9.6, you will add the costs, utilities and a discount rate to your model, before programming the calculation of the deterministic incremental cost effectiveness ratio (ICER) in Sect. 9.7.

## 9.2 The Model

First, we need to introduce you to the conceptual model that you are going to build. We have not chosen a specific disease, but rather a stylised representation of a chronic disease which is characterised by periods of controlled disease with periodic disease flairs and an accumulation of long-term disability. This type of structure would suit conditions such as rheumatoid arthritis, cardiovascular disease and diabetes. We are going to consider the simplest form of cost effectiveness analysis – a comparison of two alternative treatments in terms of their costs and outcomes.

Whilst this chapter covers the major steps in building a deterministic model in Excel, to provide additional support, we have a series of worked exercises using Excel Workbooks available from <http://medhealth.leeds.ac.uk/costeffectivenessmodelling>. A total of five exercises are provided that (a) fill in the transition matrix, (b) and (c) track a cohort of patients across time, (d) add a discount rate to the model; and (e) add a half-cycle correction. These exercises track the tasks set out in this chapter but provide a more hands-on approach.

Real-world models for any of these diseases would break out the uncontrolled disease state into multiple states, as well as having multiple disability states. However, we do not require this added complexity for the purpose of demonstrating the techniques covered in this book. Figure 9.1 shows the influence diagram for the model. For convenience, we have provided labels for the transition probabilities that will be required in the model. We will use these labels throughout the exercise, and it may be useful to use these as ‘names’ in Excel as you build the model.



**Fig. 9.1** Influence diagram for the Markov model

**Table 9.1** Parameter table

<b>Transition probabilities</b>		<b>Utility parameters</b>
Chronic silent to chronic active (p_active)		Utility in chronic silent disease (u_silent)
Chronic silent to complications (p_compl_sil)		Utility in chronic active disease (u_active)
Chronic silent to death (p_dead_sil)		Utility in complications (u_comp)
Chronic active to chronic silent (p_success)		Utility in death (u-death)
Chronic active to chronic complications (p_compl_act)		
Chronic active to death (p_dead_act)		
Complications to death (p_dead_cmp)		
<b>Cost parameters</b>		
Cost in chronic silent (c_silent)		Cost in complications (c_comp)
Cost in chronic active (c_active)		Cost in death (c_death)

**Table 9.2** Model transition matrix

		Outcome at end of period/start of next period			
		Silent	Active	Compl	Dead
Outcome at start of period	Silent	1-(p_silent+p_compl_sil+p_dead_sil)	p_silent	p_compl_sil	p_dead_sil
	Active	p_success	1 - (p_success + p_compl_act + p_dead_act)	p_compl_act	p_dead_act
	Compl	0	0	1-p_dead_cmp	p_dead_cmp
	Dead	0	0	0	1

Remember, in addition to transition probabilities, each state will need cost and utility data. Can you construct a full model parameter list (Step 1)? You might also want to identify a consistent naming structure for all parameters.

Does your parameter table look like Table 9.1? If it doesn't, what parameters did you miss and why? Did you have any parameters that we have not listed? What type of state is Dead in this model?

Once you have constructed the list of parameters that will be required for the model, you need to identify the transition matrix, which will be the model's engine (Step 2). Using the variable names in Table 9.1, construct the model transition matrix and then compare it to our version in Table 9.2.

If you have successfully produced a transition matrix that looks like Table 9.2 and you are confident you understand why it should look like that, then you are ready to start constructing the model in Excel. To help you construct the model efficiently, we are going to go through some 'good design' principles for modelling in Excel.

### 9.3 Modelling in Excel

Remarkably, the construction of a probabilistic Markov model in Excel uses relatively few functions and relies quite heavily on the use of some simple Visual Basic macros. However, those functions and macros will generate a great deal of data, and the inexperienced modeller can easily lose track of what is happening and mistakes can easily creep in and go undetected. Experience has shown that a few simple strategies can help keep control of the data and make model construction and debugging much easier.

Our first recommendation is to have a single worksheet containing all the parameters in the model. By using the naming facility within Excel and then referring to parameters by their names in the formulas, any modification of the parameter values needs only to be done once, on the parameters worksheet. The amended parameter value will then be used by all formulas in the model. This avoids the need to amend the parameter in every formula that uses it. This approach hardwires quality control into the model construction and saves huge amounts of time when changing parameter values in the light of new evidence.

Step 3 in this exercise is to open up a new Excel worksheet and construct a parameter table – reflecting the information in the parameter table you constructed in Step 1. Remember, eventually, we are going to construct a probabilistic model, so you will want to create a parameter table that has columns for the moments required to specify the parameter distribution. Think about the types of distributions we use for each type of data in order to choose which moments may need to be specified for each type of data. You will also want to name the worksheet so that any future user of the model will know that this first worksheet contains the parameter table.

Figure 9.2 shows the parameter sheet in our model. As you can see, we have rows for each of the model parameters, and for each model parameter, we have allowed space to specify the distribution parameters. Those parameters we expect to be different for the two treatments have two rows, one for each treatment-specific parameter. Did you remember to put in extra rows for the treatment-specific parameters? Our parameter table also includes a column for the deterministic parameter values and another column called ‘value used’. By doing this, we will be able to compare the results of our deterministic and probabilistic analyses. How will we tell Excel which values to use for the analysis? Hint: an ‘IF statement’ that references a ‘model-type’ cell will be useful.

The final component of Fig. 9.2 that you may not have included in your parameter table is the ‘Refers to’ column. This is included to help the user rather than the person who built the model. It simply tells the user which interventions the parameter applies to. Hence, it takes three values: Treatment A, Treatment B and Common, the latter being for parameters that are the same for both treatments.

Figure 9.3 is also taken from our parameters worksheet, and it shows a very useful quality control check. For each parameter, we record the mean and standard deviation from the data used to construct the distribution, and next to it,

**Fig. 9.2** Parameter table structure

Transition parameters	From Data		From dist	
Parameter	Mean	SD	Mean	SD
p_active				
p_success				
p_compl_sil				
p_compl_act				
p_dead_sil				
p_dead_act				
p_dead_cmp				
<b>Cost parameters</b>				
Parameter	Mean	SD	Mean	SD
Chronic, silent disease				
Chronic, active disease				
Complications				
c_TxA				
c_TxB				
<b>Utility parameters</b>				
Parameter	Mean	SD	Mean	SD
Chronic, silent disease				
Chronic, active disease				
Complications				

**Fig. 9.3** Data source and prediction quality control in parameter sheet

we have cells to capture the mean and standard deviation predicted in the model. Whilst there is likely to be (very) small differences due to sampling, if these data and the predictions are substantially different, then there is a problem with the model that must be resolved before the results can be used with any confidence. We would recommend that you add this to your parameters worksheet also.

## 9.4 Constructing the Parameter Table

Now, having created a parameter table, we need to create a worksheet that has the transition matrices. We want this to be a separate sheet, so that users who want to examine the transition matrix do not have to search for it within a worksheet that contains other information. We suggest you call this worksheet ‘transition matrix’. If you name the cells in your parameters worksheet using the same labels as you used to complete Table 9.2, then constructing your transition matrix is as simple as transposing Table 9.2 into the ‘transition matrix’ worksheet. Remember to put the ‘=’ sign before the text, so that Excel recognises them as formulas.

As well as the transition matrix, you will need to create a space for the Markov traces for each treatment to be recorded over the time horizon of the model (Step 4). Call the treatments Treatment A and Treatment B. The model we are going to build will have a 5-year time horizon and a 1-month cycle.

**a**

	I	J	K	L	M
3		Silent	Active	Complications	Dead
4	Silent	=1-K4-L4-M4	=p_active_AA	=p_compl_sil_A	=p_dead_sil_A
5	Active	=p_success_A	=1-J5-L5-M5	=p_compl_act_A	=p_dead_act_A
6	Complications	0	0	=1-M6	=p_dead_cmp_A
7	Dead	0	0	0	1

**b**

Treatment A		Probabilities					Costs					Outcomes				
Years	Period	Silent	Active	Compl	Dead	discount	Silent	Active	Compl	Dead	Silent	Active	Compl	Dead		
0.00	1															
0.01	2															
0.02	3															
0.03	4															
0.04	5															
0.05	6															

**Fig. 9.4** (a) Transition matrix and Markov trace worksheet structure. (b) Markov trace structure for treatment A

- How many rows will you need to capture the Markov trace?
- How many columns will you need to capture the Markov trace for each treatment?
- What will be the heading for each column in the Markov trace?

Remember that the total costs and outcomes for each treatment are driven by the Markov trace; your structure needs to include space to capture these data.

Figure 9.4a shows our transition matrix worksheet for Treatment A. The structure for Treatment B is identical. The Markov trace table, Fig. 9.4b, consists of three sub-tables, one each for the distribution of the cohort, the costs of care and the outcomes (QALYs). In each sub-table, we have one column for each of the four states in the model.

If you look carefully, you will notice we have also included a column for the discount rate. Adding a discount rate to your worksheet is the next step in the exercise. Make sure you add it to the trace data for both treatments.

## 9.5 Programming Your Model

Having established the structure to capture the Markov trace data, it is time to write the functions that will actually generate that data. This is Step 5 in our exercise. At the centre of a Markov model is the application of the transition matrix to the cohort. First, you need to fill in the formulas for the transition matrix. If you have named the parameter cells in your parameters worksheet, then you can simply use the names and complete the transition matrix for Treatment A using the information in Table 9.2 above. Remember, Excel will want you to use an '=' sign at the start of each calculation so that it knows you are asking it to do a calculation. Otherwise, it will think you are simply entering text.

Table 9.3 gives you the deterministic parameter values that you will need to use in your parameters worksheet in order for the transition matrix formulas to return actual values. You will need to reference these values in the ‘Value Used’ cells in the parameter sheet and give them the appropriate names for use in subsequent formulas. Here, you will need to use the ‘IF statement’ that tells Excel where to look for the data depending upon whether the model is deterministic or probabilistic. You will need to do this for the transition matrix parameters and the cost and utility parameters. We will provide the data for these parameters shortly.

Before we can generate the Markov trace for the distribution of the cohort across the four health states in the model, we need to specify the starting distribution, i.e. how the cohort of patients is distributed across the four states at the start of the model. We are going to assume that all the patients are in the active disease state at the start of the model. Therefore, in Period 1, we put the value 1.0 in the active disease column and 0.0 in the other four columns.

Step 6 is to apply the transition matrix to the cohort to create the Markov trace for all 60 periods of the model time horizon. Remember that the transition matrix and the distribution of the cohort are arrays that can be multiplied in Excel using the matrix multiplication function =MMULT. Remember to allow the cell references in the multiplication to change with each period of the model.

Table 9.4 shows the formulas we have used to generate the trace over the first five periods. Your cell references may vary if you have positioned the Markov trace differently in your worksheet. In order to interpret the formulas, note that J22:M22 contain the Period 1 probabilities for the four states (the second row of the table); J23:M23 contain the Period 2 probabilities for the four states (third row), and so on. Here, \$J\$4:\$M\$7 contains the transition matrix, as in Fig. 9.4a. Note that all the

**Table 9.3** Parameter values for the transition matrices

	Probabilities	Treatment A	Treatment B
p_active	0.094	0.093	
p_success	0.194	0.212	
p_compl_sil	0.012	0.014	
p_compl_act	0.144	0.098	
p_dead_sil	0.002	0.002	
p_dead_act	0.001	0.001	
p_dead_cmp	0.003	0.003	

**Table 9.4** Functionality for Markov trace showing distribution of cohort

Period	Probabilities			
	Silent	Active	Compl	Dead
1	0	1	0	0
2	(=MMULT(J22:M22, \$J\$4:\$M\$7))	(=MMULT(J22:M22, \$J\$4:\$M\$7))	(=MMULT(J22:M22, \$J\$4:\$M\$7))	(=MMULT(J22:M22, \$J\$4:\$M\$7))
3	(=MMULT(J23:M23, \$J\$4:\$M\$7))	(=MMULT(J23:M23, \$J\$4:\$M\$7))	(=MMULT(J23:M23, \$J\$4:\$M\$7))	(=MMULT(J23:M23, \$J\$4:\$M\$7))
4	(=MMULT(J24:M24, \$J\$4:\$M\$7))	(=MMULT(J24:M24, \$J\$4:\$M\$7))	(=MMULT(J24:M24, \$J\$4:\$M\$7))	(=MMULT(J24:M24, \$J\$4:\$M\$7))
5	(=MMULT(J25:M25, \$J\$4:\$M\$7))	(=MMULT(J25:M25, \$J\$4:\$M\$7))	(=MMULT(J25:M25, \$J\$4:\$M\$7))	(=MMULT(J25:M25, \$J\$4:\$M\$7))

formulas in Table 9.4 are surrounded by curly brackets “{ }”. This tells Excel that it should calculate the four probabilities in each row together as an array.

## 9.6 Adding a Discount Rate, Costs and Utilities

You should now have data showing the distribution of the cohort over all 60 cycles (5 years) of the model. The next task is to add in the calculation for the discount rate in the relevant column. Remember that, by convention, we apply the same discount rate to each year and the first year is not discounted. Check back to Chap. 5 (Sect. 5.6) to get the formula. The Excel symbol for raising something to a power is ‘^’. You may want to add the discount rate as a parameter in the parameter table. This will make it easier to examine whether varying the discount rate impacts upon your results. If you do add discount rate into the parameter table, remember to name the cell and use that in your functions. For the primary analysis, use a discount rate of 5 % per annum for both costs and benefits.

Once you have added in the discount rate, you need to go through the same process for Treatment B: construct the transition matrix, set the initial distribution for the cohort, construct the Markov trace for the distribution of the cohort across the states and then add in the discount rate.

Table 9.5 gives the monthly costs and utilities for each state in the model. You need to add these into your parameter table and then complete the Markov trace data

**Table 9.5** Costs and utility parameters

	Treatment A	Treatment B
<i>Cost</i>		
Silent	\$61.52	\$926.52
Active	\$122.43	\$987.43
Comp	\$397.00	\$397.00
Dead	\$0.00	\$0.00
<i>QALY parameters</i>		
Silent	0.068	0.068
Active	0.054	0.054
Comp	-0.011	-0.011
Dead	0.000	0.000

						Silent	Active	Compl	Dead	Silent	Active	Compl	Dead
Totals	1830	1107	5.81	39.42	3.70	653.18	695.25	14461.28	0.00	0.72	0.31	-0.42	0.00
								TOTAL COST	\$15,809.71			TOTAL DAYS	0.62

**Fig. 9.5** Markov trace totals

Silent	Active	Compl	Dead
=0.5*O22+SUM(O23:O81) +0.5*R82	=0.5*P22+SUM(P23:P81) +0.5*P82	=0.5*Q22+SUM(Q23:Q81)+ 0.5*Q82	=0.5*R22+SUM(R23:R81) +0.5*R82
		Total Cost	=O90+P90+Q90+R90

**Fig. 9.6** Formulas for half-cycle correction of costs

		Costs	Outcomes	Net benefit
Treatment A		\$15,609	1.184	\$19,903
Treatment B		\$19,586	1.278	\$18,740
Incremental		\$3,977		
		for	0.094	QALYs
ICER		\$42,411.02	per QALY	

**Fig. 9.7** Expected incremental cost effectiveness and net benefit (deterministic)

for costs and outcomes. If you set up the formulas for calculating the costs and benefits in the first period, you should be able to simply copy these down for all periods in the model. Once you have done this, you need to calculate the total time, costs and benefits in each state over the time horizon of the model. We recommend locating these calculations at the bottom of the Markov trace columns. Figure 9.5 shows this information for Treatment A. You will need to do this for both Treatment A and B.

## 9.7 Adding the Calculation of the Deterministic Incremental Cost Effectiveness Ratio (ICER)

At this point, you have very nearly constructed a complete deterministic Markov cost effectiveness model. All that remains is to apply the half-cycle correction to the Markov trace data. Refer back to Chap. 5 to remind yourself why we apply a half-cycle correction and then add a ‘half-cycle corrected’ set of Markov trace totals below the uncorrected totals that you have just calculated. Figure 9.6 shows the formulas we used to calculate the half-cycle corrected cost totals in our model. Your cell references may differ depending upon where you have located the Markov trace data in your worksheet, but the formulas should be the same. In our worksheet, Row 22 contains the first (half) period’s results and Row 82 contains the extra (half) period’s results used when calculating the half-cycle correction.

Now you need to calculate the ICER and the Net Benefit. We suggest you set  $\lambda = \$50,000$  per QALY. Having set our hypothetical cohort as one person, the total costs and outcomes are in fact the expected costs and outcomes, so you can use these results to calculate the ICER. Before you do that however, we suggest that you create a third worksheet: the results worksheet. Name the cells that have the expected costs and outcomes in the Markov trace worksheet, so that you can use these names in the formulas on your results worksheet. Figure 9.7 shows the results for our model.

The structure for your cost effectiveness model is now complete. In the next chapter, you will replace the deterministic parameters with probability distributions.

## 9.8 Summary

- We have outlined 6 steps in constructing a deterministic cost effectiveness model:
  - Step 1: Construct a full model parameter list.
  - Step 2: Construct the transition matrix.
  - Step 3: Construct a parameter table – reflecting the information in the parameter table you constructed in Step 1.
  - Step 4: In addition to the transition matrix constructed in Step 2, create a space for the Markov traces for each treatment to be recorded over the time horizon of the model. Include sub-tables, one each for the distribution of the cohort, the costs of care and the outcomes (QALYs). If applicable, add a discount rate to your worksheet. Make sure you add it to the trace data for all treatments.
  - Step 5: Write the functions that will generate that data. You need to fill in the formulas for the transition matrix; ensure you specify the starting distribution.
  - Step 6: Apply the transition matrix to the cohort to create the Markov trace for all periods of the model time horizon. Ensure you include the calculation for the discount rate in the relevant column and, if applicable, apply the half-cycle correction to the Markov trace data. Calculate the ICER and Net Benefit.

# Chapter 10

## Making a Markov Model Probabilistic

**Abstract** The deterministic model outlined in the previous chapter is often used as an important step towards constructing a stochastic model. Constructing the deterministic model and conducting one-way sensitivity analyses can aid the analyst in checking the validity of the underlying model. Given the importance of checking that a model is performing as expected, it is good practice to build your models with the ability to operate in either deterministic or stochastic mode.

### 10.1 Introduction

In this chapter, we will work through the process of converting the deterministic cost effectiveness model you constructed in Chap. 9 into a fully probabilistic model. Prior to this we will introduce you to some additional functions in Excel, including the construction of Excel macros to automate the repetitive components of the simulation component of the probabilistic analysis. Section 10.2 recaps some general issues around deterministic and probabilistic cost effectiveness models. Sections 10.3 and 10.4 describe how to make the model parameters stochastic in Excel and how to automate the simulation process. Sections 10.5, 10.6 and 10.7 take you through the exercises for making all effectiveness, cost and utility parameters probabilistic and incorporating the Cholesky Decomposition.

### 10.2 Deterministic and Probabilistic Cost Effectiveness Analysis

However, first we will recap some of the general issues around deterministic and probabilistic analyses. When we introduced sensitivity analysis in Chap. 4, we explained that there are many ways to explore the unavoidable uncertainty around our expectation of the costs and benefits of alternative actions. Probabilistic sensitivity analysis (PSA) represents one of the most powerful tools available in

economic evaluation to capture uncertainty over a large number of different pieces of evidence.

It would be misleading to suggest that PSA allows ‘everything’ to be uncertain, since this is clearly not the case. Rather, a PSA is useful to explore the uncertainty in the values taken by a set of parameters in a model. It is not appropriate to examine the core assumptions of a model such as the clinical pathway. For example, in a model considering the role of how to manage suspected strokes, a decision tree might consider the question of how we identify potential cases, how we manage cases that are identified (as well as those that are not) and the consequences of various complications that may arise due to management. A stochastic model – the type of model that allows us to ‘do’ a PSA – would incorporate these structural issues in exactly the same way. However, the parameter distributions would seek to capture all uncertainty related to sampling variation, extrapolation and generalisability. Within the stroke example identified above, the parameters might deal with issues such as the population incidence of stroke, the sensitivity of a particular method to detect stroke or the costs and quality of life lost attached to a particular outcome.

As a broad rule, if a deterministic model coded in Excel would use a number, then it is a potential candidate to be varied within a stochastic model. As a corollary, those things are not incorporated using numbers that are more likely to be varied in other types of sensitivity analyses such as those assessing structural and methodological uncertainty. However, this is not an absolute rule. For example, key methodological choices, such as the discount rate, are numerical parameters in a model but are not normally included in a PSA. Methodological uncertainty of this type would typically be assessed, and it is normally done by changing the value of and rerunning the entire PSA for this new value. Hence, three types of parameters are typically varied within a full PSA: effectiveness, costs and utilities. We have reviewed the distributions that are typically used for each type of parameter in Chaps. 6, 7 and 8.

Before moving on to the details of how to make a model probabilistic, it is worth taking a moment to recognise the value of constructing a deterministic model. The deterministic model is often used as an important step towards constructing a stochastic model, since the problems inherent in one-way sensitivity analyses can aid the analyst in checking the validity of the underlying model. In one-way sensitivity analyses, we essentially make a simplistic assumption that we can consider each item individually – but doing so is important when we check that the model is working as it is meant to. For example, if we were to increase an effectiveness parameter that determines the mortality rate and find that life expectancy *increases*, then we have a *prima facie* case that there is an error in the model. The fact that everything else remains constant when we change a parameter means that we are able to quickly check that things work in the way that they are meant to. In contrast, a model built to be stochastic from the beginning does not allow us to do this.

A perfectly functioning stochastic model will, other things being equal, have lower life expectancy if the mortality rate increases. Other things are not generally equal though, and if changes in the parameters other than mortality rate increase the life expectancy by a large enough amount, we would expect that the life expectancy could appear *higher* if we manually entered a lower mortality rate. So, a stochastic

model doesn't allow us to check the face validity (i.e. *Does it 'look' right?*) of the model as reliably as the deterministic model. Given the importance of checking that a model is performing as expected, it is good practice to build your models with the ability to operate in either deterministic or stochastic mode.

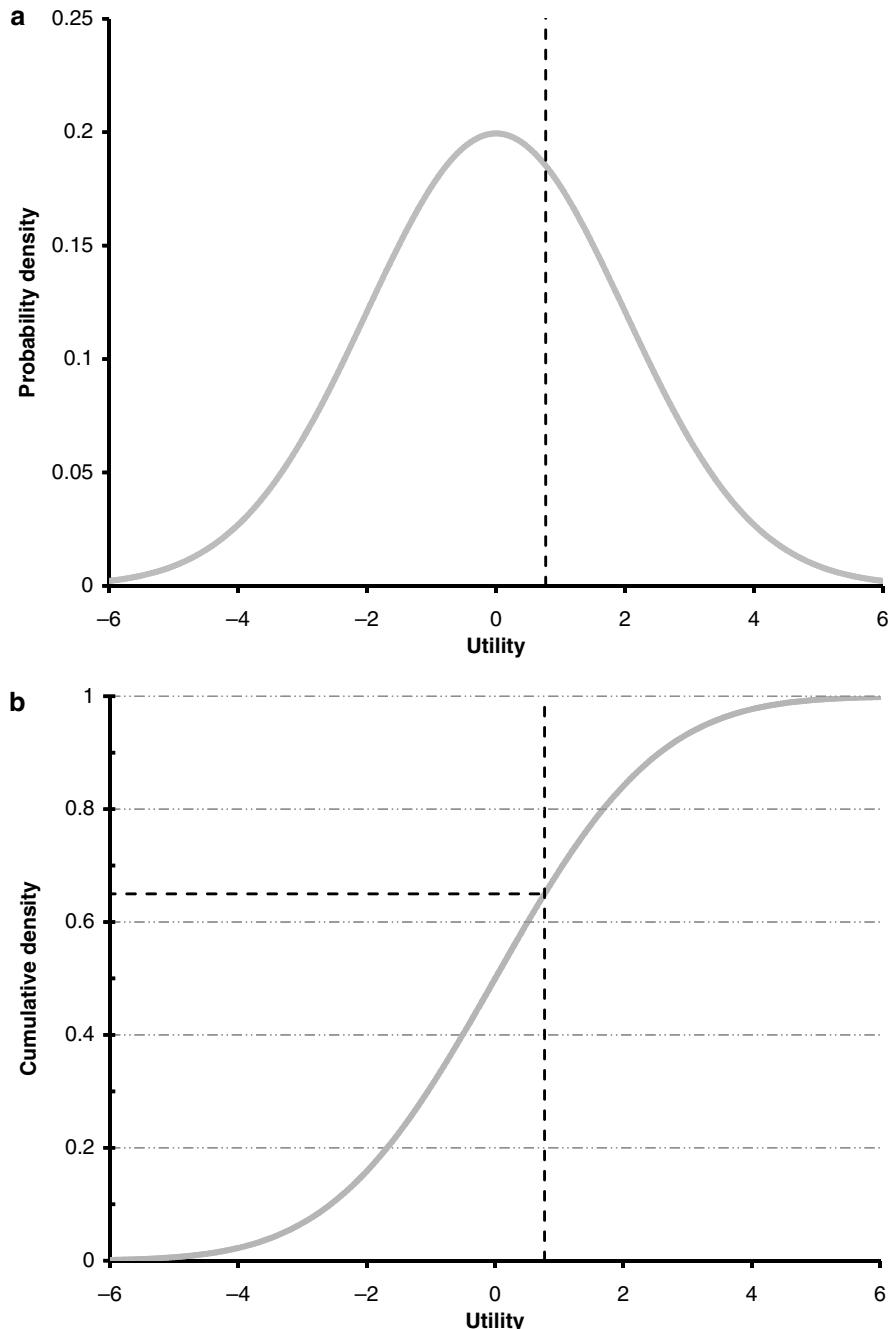
## 10.3 Making Model Parameters Stochastic

Obtaining values for a stochastic parameter uses a relatively standard type of formula. As we are looking to find a value that is drawn from a distribution at random, the key piece of information is a random number. In Excel, this is identified by the RAND() function, which randomly draws an uniformly distributed real number between 0 and 1. This function is re-evaluated every time a worksheet is calculated, either because something changes in the spreadsheet (e.g. a new formula is entered or something is pasted) or because the user presses the F9 key (in Windows) to tell Excel to recalculate. This makes it ideally suited for a stochastic model, in which these types of recalculations are necessary.

We use the RAND() to identify a point within a distribution that we are interested in sampling from. For example, suppose that we wish to sample from a Normal distribution with mean of 0 and a standard deviation of 2. The upper portion of Fig. 10.1 depicts the central portion of this distribution between the values of -6 and +6; although both bigger and smaller values are possible outside of this range, it is pretty clear that the majority of random draws from this distribution will fall within this range. The lower portion of Fig. 10.1 shows the same information, but as the cumulative density function (CDF). As an illustration, suppose that we drew a random number of 0.65. We might then ask what value we would need to pick so that the number we picked was higher than 65 % of values. Using the random number in this way means that we are equally likely to choose any point 'along' the distribution (in terms of the cumulative probability), but are much more likely to pick numbers that are more common (have higher density) within the distribution.

Excel has a relatively wide range of built-in functions that capture cumulative density functions (see Table 10.1). In the example above, we would use the following formula; =NORM.DIST(A1,0,2,TRUE) - where cell A1 would contain the random number draw, using =RAND(). This is read by Excel as follows – the Normal distribution is evaluated at the value in Cell A1, where the specific distribution used uses a mean of 0 and a standard deviation of 2. Finally the 'TRUE' component confirms to Excel that we want to evaluate the cumulative density function (i.e. the lower portion of Fig. 10.1).

Fortunately, each of the distributions in Excel also has a corresponding inverse function (again, see Table 10.1), which allows us to ask at what value the cumulative density function equals a given figure. This, then, provides us with an easy way of finding the relevant numbers in a probability draw. For the example of the Normal distribution with a mean of 0 and a standard deviation of 2, we could ask Excel to draw a number as NORM.INV(RAND(), 0, 2). This also means that we can draw probabilities without having to search for results (Excel does this for us).



**Fig. 10.1** Truncated normal distribution  $(0,2)$

**Table 10.1** Frequently used Excel supported distributions, functions and inverse functions

Distribution	Distribution function	Inverse distribution function
Normal	NORM.DIST	NORM.INV
Binomial	BINOM.DIST	BINOM.INV
Beta	BETA.DIST	BETA.INV
Gamma	GAMMA.DIST	GAMMA.INV
LogNormal	LOGNORM.DIST	LOGNORM.INV
<i>t</i> -distribution	T.DIST	T.INV.2T
Weibull	WEIBULL.DIST	

More generally, each of the families of parameter distributions found in Table 10.1 will have its own set of associated parameters (or ‘parameterisation’). In Excel, when we draw values from this distribution, there is a similar structure, so that our command will typically begin with the relevant inverse distribution function and choose a random number (or cell containing a random number) and then the parameterisation of the distribution that we want to use. It is important that each parameter distribution references its own random number draw [=RAND()]. If each distribution used the same random number draw this would create an artificial correlation between parameters.

## 10.4 Obtaining a Probabilistic Sensitivity Analysis from a Stochastic Model

Before we move on to programming the PSA, it is important to make sure that the outputs from the analysis are grouped together. In the previous chapter, you set up cells to capture the single cost and outcome estimates for each of the treatments. Now you need to establish cells to capture and store the outputs from each simulation, for later analysis. This will entail setting up a set of cells that are equal to each of the outcome cells you created at the end of the previous chapter. But it will also require that at the end of each simulation, these results are copied into a results storage area. If you are running 10,000 simulations, then you will want to automate this in order to get the job done before you retire!

Whilst this chapter covers the major steps in making your model probabilistic, to provide additional support, we have a series of worked exercises using Excel Workbooks available from <http://medhealth.leeds.ac.uk/costeffectivenessmodelling>. A total of six exercises are provided that (a) modify your deterministic model to accept stochastic parameters; (b) and (c) add Beta and approximated Dirichlet distributions; (d) add Gamma and Log Normal Distributions for costs; (e) add stochastic utility parameters and (f) add correlated parameters using Cholesky Decomposition.

The way to automate repetitive activities in Excel is to create a macro. A macro is simply a Visual Basic programme. If you have never written a macro before, then the easiest way is to use the record macro function and ‘walk’ Excel through the operations that you wish to automate. Excel will record the Visual Basic code and then let you edit this to control how many times the action should be repeated.

In Box 10.1 we have produced example code for running 10,000 simulations for a PSA. You can see we have named the results cells ‘psa\_source’ and instructed Excel to ‘copy’ these cells. We have then instructed Excel to go to the cell range named ‘psa\_target’, move down one cell and then paste the values that it just copied from ‘psa\_source’.

As long as you identify the number of individual rows you need (as NumRuns in Line 3), this will copy the output in psa\_source to this number of rows following psa\_target. You don’t need to be an expert in Visual Basic to construct a macro of this sort. However, it is worth noting here that the macro switches screen updating off and on (Lines 5, 21) and the stochastic version of the model on and off (Lines 6–7, 19–20). We only have to select and copy psa\_source once (Lines 9–10), find psa\_target (Line 12) and move down (Line 14) before pasting the values of this range NumRuns times (Lines 15–16). Excel updates the stochastic model every time the model pastes in the previous set of value, and hence there is a new set of simulation results ready to be recorded in the psa\_target range. This macro is produced using Excel’s record macro facility and is not particularly efficient. Appendix 10.1 provide more efficient macro codes for the functions used in this chapter.

You may be wondering why we chose 10,000 simulations for our macro. How should we choose how many ‘NumRuns’ to specify in the macro; how do we identify when our PSA has run ‘enough’ that we can rely on the model? The key concepts here are of reliability and convergence. We have to be confident that the PSA is reliable in the sense that it will not produce a substantially different distribution if we reran the model. The simplest way of doing this would be to rerun the model for the base case

### Box 10.1: Sample Macro for Probabilistic Sensitivity Analysis

```

1 Sub PSAruns ()
2
3     Dim NumRuns As Integer = 10000
4
5     Application.ScreenUpdating = False
6     Application.Goto Reference:="opt_modeltype"
7     ActiveCell.FormulaR1C1 = "Stochastic"
8
9     Application.Goto Reference:="psa_source"
10    Selection.Copy
11
12    Application.Goto Reference:="psa_target"
13    For counter = 1 To NumRuns
14        ActiveCell.Offset(1, 0).Select
15        Selection.PasteSpecial Paste:=xlPasteValues, _
16        Operation:=xlNone, SkipBlanks:=False, Transpose:=False
17    Next
18
19    Application.Goto Reference:="opt_modeltype"
20    ActiveCell.FormulaR1C1 = "Deterministic"
21    Application.ScreenUpdating = True
22 End Sub

```

(i.e. for the case without modifying your methodological assumptions) and check to see if the mean and standard deviations of each cost and effectiveness distribution are the same (to an acceptable number of significant figures). If these numbers are the same, then we can say that the results have converged and the probabilistic sensitivity analysis outputs can be safely used. Note, if you change any of the methodological assumptions, then strictly you should re-establish convergence rather than assume that the same number of simulations will be sufficient. As a rule it is helpful to start most models with 5–10 thousand runs, since this usually takes relatively little time and gives a good starting point. If the PSA results do not converge, a higher number of runs will be required. Some element of trial and error is normally required to identify what is ‘enough’.

## 10.5 Exercise: Probabilistic Effectiveness Parameters

Using the deterministic cost effectiveness model you constructed in Chap. 9, the first task in making the model probabilistic is to tell Excel to change the values used from the deterministic values to the probabilistic values. This is done in two steps – first change the model type cell from ‘deterministic’ to ‘stochastic’ (in our model, opt\_modeltype). Then, in the ‘Values used’ column on the parameter worksheet, make sure that the IF-statement now directs Excel to use the data in the ‘Stochastic’ column of the parameter table rather than the ‘Deterministic’ column. You will need to do this for each parameter.

Having set up the “Values used” column in the parameter worksheet to pick up the stochastic data, you now need to ensure that the cells in the transition matrix pick up the data from the ‘Values used’ column. In our model, the transition matrix is in cells B22:C28 on the transition matrix worksheet, and the ‘Values used’ cells are in D8:D20.

We are going to use a Beta distribution to characterise the effectiveness parameters in our model. On the parameter worksheet the relevant columns are:

- RAND=random number used for the simulation
- Distribution=type of distribution used (Beta, Gamma, LogNormal, etc...)
- Param\_1=value for the first parameter of the distribution
- Param\_2=value for the second parameter of the distribution
- Data moment\_1=data estimate of the first moment of the distribution
- Data moment\_2=data estimate of the second moment of the distribution
- Distribution moment\_1=moment 1 from the distribution
- Distribution moment\_2=moment 2 from the distribution

The following data is provided on the effectiveness of Treatment A.

Transitions from SILENT: 125 people were followed up for a month.

113 people remained in SILENT.

11 entered ACTIVE.

1 entered COMPL.

0 entered DEAD.

Of the 125 people:

How many people went from Silent disease to Active? \_\_\_\_\_

How many people did not go from Silent to Active? \_\_\_\_\_

Use this data to characterise a Beta distribution for  $p_{\text{active}}$  (the probability of moving from Silent to Active disease). You will need to programme a random number draw in the cell in the RAND column. If necessary check back to Sect. 6.5 to remind yourself of the random draw function in Excel. Remember, for the proportions version of the Beta distribution, the number of people who progressed is the value for the alpha parameter, and the number of people who didn't progress is the beta parameter. These data go in the Param\_1 and Param\_2 columns. Now you need to use the Beta.inv() function in the stochastic column to create a random draw from each of the distributions.

Using the formulas below calculate the mean and variance produced by the distribution you have constructed, and record these in the Distribution moment\_1 and Distribution moment\_2 columns, respectively. Calculating these figures serves as a check, especially where we have used the method of moments to generate the parameters in a distribution.

$$\mu = \frac{\alpha}{\alpha + \beta}$$

$$\sigma^2 = \frac{\alpha\beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)}$$

In Table 10.2 we report the observed transitions from a 1-month follow-up for Treatment A and Treatment B. Use these data to parameterise beta distributions for the remaining transition parameters for both treatments. You will note that the sample size for Treatment B is much larger than for Treatment A. What impact do you expect this to have on the standard deviations for the parameters for Treatment B compared to Treatment A? Remember that the standard deviation is the square root of the variance.

**Table 10.2** Observed 1-month transitions for Treatments A and B

Treatment A	Silent	Active	Complications	Dead
<i>Silent</i>	113	11	1	0
<i>Active</i>	12	44	8	1
<i>Complications</i>			1,426	5
<i>Dead</i>				
Treatment B	Silent	Active	Complications	Dead
<i>Silent</i>	1,010	105	16	2
<i>Active</i>	178	579	82	1
<i>Complications</i>			1,426	5
<i>Dead</i>				

	D	F	G	H
49	random draw	alpha gamma draw		parameters
Probability of <b>silent disease</b>	=RAND()	=GAMMA.INV(D50, F50, 1)	=G50/SUM(\$G\$50:\$G\$52)	
50 from silent disease				
Probability of <b>active disease</b>				
51 from silent disease	=RAND()	113 =GAMMA.INV(D51, F51, 1)	=G51/SUM(\$G\$50:\$G\$52)	
Probability of <b>complications</b>				
52 from silent disease	=RAND()	1 =GAMMA.INV(D52, F52, 1)	=G52/SUM(\$G\$50:\$G\$52)	
Probability of <b>death</b> from				
53 chronic silent disease	=RAND()	0		

**Fig. 10.2** Constructing the gamma distributions for a Dirichlet

Of course, modelling each of the individual transitions as Beta distributions ignores the inherent correlation between the transition rates from each starting point. Using a Dirichlet distribution allows us to characterise the uncertainty in the transitions whilst recognising these correlations.

Look back to Chap. 6 to refresh your memory on constructing the Dirichlet distribution using Gamma distributions. Note that each parameter will need a row for each Gamma distribution that is combined in the Dirichlet, plus a row for combining them. You will need to expand your parameter table for each of the transition parameters, and then link transition matrix to the Dirichlet draws in the ‘Values used’ column. Remember that the Dirichlet distribution for each transition parameter from a specific health state is obtained by dividing gamma draw for the transition parameter by the sum of the gamma draws for all of the transition parameters from that state.

To help you, Fig. 10.2 reproduces the Excel formulas used to construct the Dirichlet for Treatment A from the Silent State to Active Disease (p\_active), Complications (p\_compl\_sil) and Dead (p\_dead) from our version of the model.

## 10.6 Exercise: Probabilistic Cost and Utility Parameters

Table 10.3 reports the data required to parameterise the cost parameters for each state in the model, except Dead. We assume that dead patients do not incur any further costs. It also specifies the distribution that you need to parameterise. If you look back to Chap. 7, you can refresh your memory on how to parameterise each of these distributions.

Having added Rand() in the RAND column for the Silent and Active cost parameter rows, first calculate the standard deviation for the LogNormal distribution using the following formula:

$$\sigma = \sqrt{\ln\left(1 + \frac{s^2}{E[x]^2}\right)}$$

**Table 10.3** Cost data to parameterise cost distributions

State	Type	Mean	Standard deviation
SILENT	LogNormal	\$28.21	\$3.98
ACTIVE	Lognormal	\$84.20	\$22.10
COMPL	Gamma	\$432.69	\$18.01

Now calculate the mean for the distribution using the formula

$$\mu = \ln(E[x]) - \frac{1}{2}\sigma^2$$

Enter these parameters into Param\_1 and Param\_2 columns for the Silent and Active states, respectively. Then draw a value from the distribution using the LOGNORM. INV function, in the stochastic column. Now use the following formulas to enter the quality assurance data checks in the Distribution moment columns.

$$E[x] = e^{\mu + \frac{1}{2}\sigma^2}$$

$$s = e^{\mu + \frac{1}{2}\sigma^2} \sqrt{(e^{\sigma^2} - 1)}$$

The cost parameter for the third state – Complications – is parameterised using the Gamma distribution. The Gamma distribution is parameterised using alpha and beta, where

$$\beta = \frac{s^2}{E[x]}$$

and

$$\alpha = \frac{E[x]}{\beta}$$

You can check that the values provided by the distribution are the same as the values used to specify the distribution using the following formulas in the Distribution moment columns:

$$E(x) = \alpha\beta$$

and

$$\sigma = \sqrt{\alpha\beta^2}$$

**Table 10.4** Utility data to parameterise utility distributions

State	Type	Mean	Standard deviation
SILENT	Beta	0.819	0.021
ACTIVE	Beta	0.653	0.109
COMPL	Gamma	-0.137	0.18

Table 10.4 provides the data required to specify the distributions for the utility parameters in the model. We provide data for the states Silent, Active, and Complications. The utility for the state Dead is zero (0) by definition.

The utility for the Silent and Active states is almost certain to be strictly positive, and therefore you can use the Beta distribution to parameterise them. By contrast, there is a significant possibility that the utility in the Complications state is less than zero, and therefore, the Gamma distribution is appropriate.

You will need to use the continuous specification of the Beta distribution. You must use the following two formulas to calculate alpha and beta from the mean and standard deviation that we have provided in Table 10.4. Remember that the standard deviation is the square root of the variance.

$$\alpha + \beta = \frac{E[x](1 - E[x])}{s^2} - 1$$

and

$$\alpha = E[x](\alpha + \beta)$$

Once you have done this, you can calculate draws for the Beta distributions for the Silent and Active utility parameters in the stochastic column of the relevant rows in your parameter table. Also, remember to do the quality assurance check to compare the data used to parameterise the distributions and the predicted moments from the distribution.

The utility for the Complications state is expected to include some negative values, hence the use of the Gamma distribution. As we said in the book, it is often useful to specify the utility parameter in disutility space to avoid illogical values, and this offer an opportunity to apply that approach. You may want to add another row to the utility component of the parameter table and use this to enter the calculations in disutility space. Beta is calculated as the standard deviation divided by the expected value, and alpha is simply the expected value divided by beta. So calculate alpha and beta using the disutility versions of the mean and standard deviation reported in Table 10.3, and then use these to construct a random draw from the Gamma distribution in the stochastic column of your parameter table. The stochastic value for the utility parameter for the Complications state is simply one minus this random draw. The ‘Value used’ cell should link to utility value, when the model type is set to ‘stochastic’.

## 10.7 Exercise: Incorporating the Cholesky Decomposition

All parameters in your model should now have distributions associated with them and it is possible to derive simulations from the model, with the proviso that there is an implicit assumption that the parameters are independent of each other, except where the interdependence is recognised in the choice of distribution – such as the use of the Dirichlet distribution for the transition probabilities. The final step in parameterising your model will be to apply the Cholesky Decomposition, to capture additional correlation between parameters – in this case, the correlation between the utility parameters in the model. You may wish to revisit Chap. 8 to refresh your understanding of the Cholesky Decomposition before proceeding with this stage of the exercise.

Table 10.5 reports the variance-covariance matrix for the regression analysis that produced the utility parameter estimates we have used in this model. In a matrix like this, the covariance between any two parameters can be found by looking at the column for one of the parameters and the row for the other one. Where we look at a value along the main diagonal (i.e. in the line from the top left to the bottom right), this gives us variances (which are the standard deviation squared).

You will need to add this to your parameters worksheet. The first step in implementing the decomposition is to calculate the standard deviations for each of these variances; i.e. the square root of the figures in the main diagonal. For example, for the ‘Silent’ state, the variance can be found where the ‘Silent’ row and ‘Silent’ column intersect (0.0004), with the standard deviation (0.021) calculated as the positive square root of this value. For the ‘Active’ state, the standard deviation equals 0.109 (i.e.  $\sqrt{0.0119}$ ). We can identify the *correlation* between the ‘Silent’ and ‘Active’ parameters by dividing the covariance (0.0009) by the standard deviations of the ‘Silent’ and ‘Active’ parameters.

Hence, the correlation between Silent and Active is calculated in the following three stages:

$$1. \sigma_{\text{Silent}} = \sqrt{0.0004} = 0.021$$

$$2. \sigma_{\text{Active}} = \sqrt{0.0119} = 0.109$$

$$3. \rho_{\text{Silent, Active}} = \frac{\text{Covariance}(\text{Silent, Active})}{\sigma_{\text{Silent}} \bullet \sigma_{\text{Active}}} = \frac{0.0009}{0.021 \times 0.109} = 0.403$$

**Table 10.5** Variance-covariance matrix for utility regression model

Variance-covariance matrix			
	Silent	Active	Compl
Silent	0.0004	0.0009	0.0008
Active	0.0009	0.0119	0.0132
Comp	0.0008	0.0132	0.0324

**Table 10.6** Correlations between utility parameters

	Silent	Active	Compl
Silent	1.000		
Active	0.403	1.000	
Compl	0.219	0.673	1.000

$\sigma_{\text{silent}}$	0	0
$a$	$b$	0
$c$	$d$	$e$

**Fig. 10.3** Components of the Cholesky decomposition for 3 correlated parameters

Now you need to calculate the correlations between each of the utility parameters, using the information in the variance-covariance matrix. Once you have completed this, your table of correlations should match Table 10.6.

Now to construct the Cholesky Decomposition for three correlated parameters, you need to calculate the following matrix of variables (Fig. 10.3).

Where each component is calculated in turn as:

$$a = \rho_{\text{Silent}, \text{Active}} \bullet \sigma_{\text{Active}}$$

$$b = \sqrt{\sigma_{\text{Active}}^2 - a^2}$$

$$c = \rho_{\text{Silent}, \text{Compl}} \bullet \sigma_{\text{Compl}}$$

$$d = \frac{\rho_{\text{Active}, \text{Compl}} \bullet \sigma_{\text{Active}} \bullet \sigma_{\text{Compl}} - ac}{b}$$

$$e = \sqrt{\sigma_{\text{Compl}}^2 - c^2 - d^2}$$

Note that a set of ‘solutions’ of this type for 2–7 variables is given in Appendix 10.1.

Having constructed the Cholesky Decomposition, the utility values in the model are constructed by combining this with (a) the observed (deterministic) mean values; and (b) some standard normal variates, as shown in Fig. 10.4.

The three cells for the stochastic utility parameters in your parameter table are now an array. Remember to use the curly brackets around your formula (press Ctrl-Shift-Enter when confirming your formula) and the MMULT function to combine the Cholesky Decomposition and the standard normal variates. The formula in our model looks like this:

$$\{= E15 : E17 + \text{MMULT}(D56 : F58, D62 : D64)\}$$

$$\begin{bmatrix} x_{\text{Silent}} \\ x_{\text{Active}} \\ x_{\text{Compl}} \end{bmatrix} = \begin{bmatrix} \mu_{\text{Silent}} \\ \mu_{\text{Active}} \\ \mu_{\text{Compl}} \end{bmatrix} + \begin{bmatrix} \sigma_{\text{Silent}} & 0 & 0 \\ a & b & 0 \\ c & d & e \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$$

**Fig. 10.4** Calculating the correlated utility draws using the Cholesky Decomposition

The cell references in your model may be different, but otherwise your formula should be the same.

The final step to generate a full set of simulation results from the Markov model is to construct a macro to repeatedly sample from the parameter distributions and record the model outputs.

Currently on the ‘Transition matrix’ worksheet, you have cells defined for calculating the expected costs and outcomes and the associated incremental cost effectiveness ratios for a single run of the model. To capture the results of multiple runs, you will need to identify a location to store the results of each individual run, before the next simulation runs. Then you will have to construct a macro that copies the values in the total cost and outcome cells and pastes them into the results storage location. It will need to do this repeatedly, for as many simulations as you choose to run. See Box 10.2 for the Visual Basic syntax of the macro in our model. Remember, each time you paste the results from one simulation, Excel will automatically recalculate the model and thus produce the next set of simulation results.

You will note that our model is set to run for 10,000 simulations. This is a reasonable number of simulations to be confident of achieving stable estimates of the expected costs and outcomes. However, if your computer is not particularly powerful, you may wish to reduce the number of simulations in order to obtain results without waiting for a few days. What would you do to establish the convergence of your model outputs?

Once your simulation has finished running, you are in a position to generate the full range of outputs that probabilistic cost effectiveness models can produce. This is the focus of Chap. 11.

### Box 10.2 Sample macro for running PSA and sorting the results for each simulation

```
Sub PSAruns()
    numruns = 10000
    'PSAruns Macro
    Application.ScreenUpdating = False
    Application.Goto Reference:="opt_modeltype"
    ActiveCell.FormulaR1C1 = "Stochastic"
```

```
Application.Goto Reference:="psa_source"
Selection.Copy

Application.Goto Reference:="psa_target"

For counter = 1 To numruns
    ActiveCell.Offset(1, 0). PasteSpecial Paste:=xlPasteValues, _
        Operation:=xlNone, SkipBlanks:=False, Transpose:=False
    Application.StatusBar = "% Complete:" & (counter / numruns)*100
Next

Application.Goto Reference:="opt_modeltype"
ActiveCell.FormulaR1C1 = "Deterministic"

Application.ScreenUpdating = False

Application.Goto Reference:="psa_results"
Selection.Copy

Application.Goto Reference:="psa_res_target"
Selection.PasteSpecial Paste:=xlPasteValues, _
    Operation:=xlNone, SkipBlanks:=False, Transpose:=False
End sub
```

## 10.8 Summary

- Probabilistic models replace single parameter values with probability distributions.
- Probabilistic sensitivity analysis addresses parameter uncertainty only. Other categories of uncertainty should not be ignored.
- In Excel, we combine the Rand() function with specific distributions functions, such as NormInv to sample values from a probability density function.
- Writing macros in Visual Basic allows us to automate the repeated sampling processes required to generate thousands of model runs required for probabilistic sensitivity analysis.
- In our exercise model, we use a construct a Dirichlet distribution from three Gamma distributions to characterise the uncertainty in the effectiveness parameters.
- We use LogNormal and Gamma distributions to characterise the uncertainty in the cost parameters and Beta and Gamma distributions to characterise the uncertainty in utility parameters.
- We applied the Cholesky Decomposition to the utility parameters to account for the correlation in the utility values for each health state in the model.

## Appendix

### *Appendix 10.1: Optimising Visual Basic Macros in Excel*

The example given in the text is designed to step a novice user through the content of a basic macro. This macro (or something like it) is based on the text that you would get if you ask Excel to record a small series of cut and paste tasks, and then insert commands to tell Excel to repeat this process.

It is possible, and recommended, that you optimise your macros as much as possible. One area where things can go wrong is when we use cutting and pasting. When Excel has to use the operating system's clipboard, it means that it becomes hard to do other normal tasks whilst your model runs 'in the background'.

For this reason, it can be easier to make greater use of names in Excel. If we use names that refer to the cells with our results, and other names to refer to the cells into which we want to copy these results, we can avoid using the clipboard.

Our original macro appears below:

#### **Box A.1: Sample Macro for Probabilistic Sensitivity Analysis**

```

1 Sub PSAruns ()
2
3     Dim NumRuns As Integer = 10000
4
5     Application.ScreenUpdating = False
6     Application.Goto Reference:="opt_modeltype"
7     ActiveCell.FormulaR1C1 = "Stochastic"
8
9     Application.Goto Reference:="psa_source"
10    Selection.Copy
11
12    Application.Goto Reference:="psa_target"
13    For counter = 1 to NumRuns
14        ActiveCell.Offset(1, 0).Select
15        Selection.PasteSpecial Paste:=xlPasteValues, _
16        Operation:=xlNone, SkipBlanks:=False, Transpose:=False
17    Next
18
19    Application.Goto Reference:="opt_modeltype"
20    ActiveCell.FormulaR1C1 = "Deterministic"
21    Application.ScreenUpdating = True
22 End Sub

```

Lines 6–7 change the value of the cell referred to as “opt\_modeltype” to the value “Stochastic”, whilst Lines 19–20 change it back to “Deterministic”. Another way of doing this would be to refer to the formula attached to the name directly, with the following used at Lines 6 and 19:

```
Range("opt_modeltype").Formula = "Stochastic"
Range("opt_modeltype").Formula = "Deterministic"
```

In Lines 9–17, the macro copies a range that includes our results and then pastes this 10,000 times. If we want to do this a little quicker, we could avoid copying the initial results. In Visual Basic, we can refer to a range relative to a cell using the Offset property. For instance, three rows down, and two rows across would mean identifying our initial cell and stating “.Offset(3, 2)”. In the model, the results in run ‘counter’ of the macro will be entered ‘counter’ rows down (and no rows across) from “psa\_target”. So here, we are interested in:

```
Range("psa_target").Offset(counter, 0)
```

As we want to change the value in these cells, we add “.Value” to the end of this and refer to the values contained in cells referenced by “psa\_source”. So, we can replace Lines 9–10 and 14–16 as:

```
Range("psa_target").Offset(counter, 0).Value = Range("psa_source").Value
```

If we make our changes, the macro is a little shorter and quite a bit quicker (Box 10.2). Note that Line 9 here continues over two lines, with “&\_” indicating to Visual Basic to read both Line 9 and Line 10 as a single instruction.

#### **Box A.2: Sample Macro for Probabilistic Sensitivity Analysis: Improved Efficiency**

```

1 Sub PSAruns ()
2
3     Dim NumRuns As Integer = 10000
4
5     Application.ScreenUpdating = False
6     Range("opt_modeltype").Formula = "Stochastic"
7
8     For counter = 1 to NumRuns
9         Range("psa_target").Offset(counter, 0).Value = &_
10            Range("psa_source").Value
11     Next
12
13     Range("opt_modeltype").Formula = "Deterministic"
14     Application.ScreenUpdating = True
15
16 End Sub
```

When we run macros, we normally turn off screen updating in Excel. This means that the computer does not spend resources refreshing the screen that could be spent in calculation. However, it can be comforting for a user to know that the computer is ‘busy’ rather than ‘stuck’, and the status bar (at the bottom of the Excel window) can be used for this purpose.

As we do not normally need the macro to update the status every period, it is worth deciding how often you want this to happen. This might be every 100 iterations (i.e. each 1 %), as in Box 10.3 below. The commands in Lines 12–15 establish whether the current value of counter is perfectly divisible by 100 (i.e. having a remainder or “modulus” of 0) and, if so, setting the value of the Status Bar to display progress in percentage terms. Having a status update of this type also gives a user a better indication as to when the model is likely to finish running.

### Box A.3: Sample Macro for Probabilistic Sensitivity Analysis: Feedback in Status Bar

```

1 Sub PSAruns ()
2
3     Dim NumRuns As Integer = 10000
4
5     Application.ScreenUpdating = False
6     Range("opt_modeltype").Formula = "Stochastic"
7
8     For counter = 1 to NumRuns
9         Range("psa_target").Offset(counter, 0).Value = &_
10            Range("psa_source").Value
11
12        If (counter Mod 100 = 0) Then
13            Application.StatusBar = &_
14            "% complete: " & (counter/NumRuns)*100
15        End If
16
17    Next
18
19    Range("opt_modeltype").Formula = "Deterministic"
20    Application.ScreenUpdating = True
21
22 End Sub

```

# Chapter 11

## Outputs from Probabilistic Sensitivity Analysis

**Abstract** In this chapter we will introduce you to the key mechanisms for summarizing the uncertainty in the outputs from cost effectiveness models; how to construct them and their strengths and weaknesses. Focus lies on probabilistic sensitivity analysis (PSA). Use of PSA ensures that the estimate of the Incremental Cost Effectiveness Ratio (ICER), and by extension the Net Benefit, is unbiased if the model is non-linear. PSA also enables us to characterise the uncertainty in the inputs to the model and thereby quantify the resulting uncertainty in the model outputs.

### 11.1 Introduction

In Chap. 4 we described how to calculate expected incremental cost effectiveness ratios (ICERs) using the outputs from a probabilistic sensitivity analysis (PSA). You will remember how we emphasised that the ICER is calculated using the means of the distributions of the costs and effects for the technologies being compared. We also explored how characteristics of a ratio make it quite difficult to characterise the uncertainty around the estimated ICER and introduced the Net Benefit Framework, which converts ICERs into a single index of expected benefit, using the decision maker's value of health, sometimes called the willingness to pay (WTP) for health or the cost effectiveness threshold, and usually denoted by the Greek letter  $\lambda$  (lambda). The formulas for the ICER, Expected Net Monetary Benefit (NMB) and Expected Net Health Benefit (NHB) (Stinnett and Mullahy 1998) are reproduced in Box 11.1.

There are two reasons for using PSA: the first is to ensure that the estimate of the ICER, and by extension the Net Benefit, is unbiased if the model is non-linear. The second is to characterise the uncertainty in the inputs to the model and thereby quantify the resulting uncertainty in the model outputs. The ICER and the Expected Net Benefit do not provide decision makers with information on the uncertainty in the model outputs, even though the PSA on which they are based has generated hundreds of thousands, if not millions of data points for exactly that purpose.

**Box 11.1: Formulas for Outputs from Cost Effectiveness Analysis**

$$\text{ICER} = \frac{\bar{C}_2 - \bar{C}_1}{\bar{E}_2 - \bar{E}_1} = \frac{\Delta C}{\Delta E}$$

$\bar{C}_2$  is the cost under the intervention of interest (the new intervention).

$\bar{E}_2$  is the effectiveness under the intervention of interest (the new intervention).

$\bar{C}_1$  is the cost under the comparator or control.

$\bar{E}_1$  is the effectiveness under the comparator or control.

$\lambda$  is the cost effectiveness threshold/willingness to pay for health.

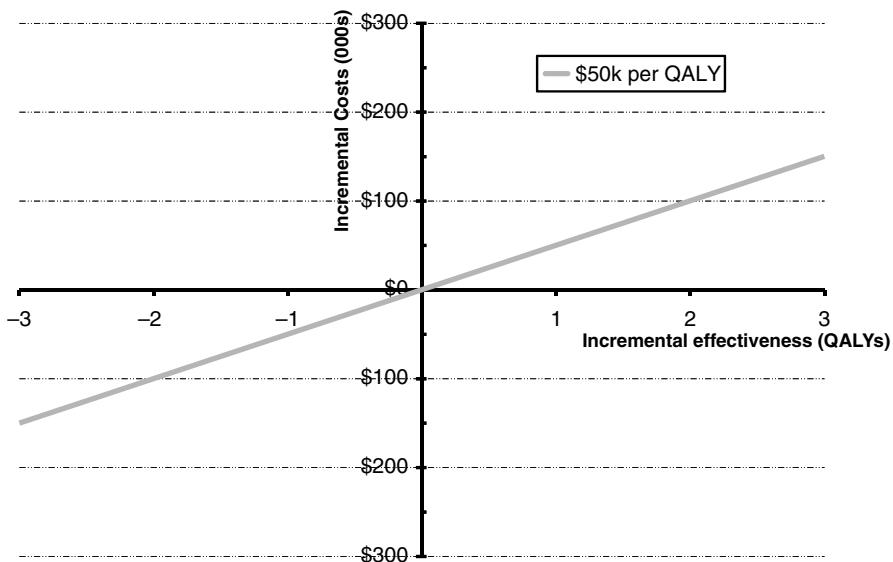
$$\overline{\text{NHB}}_i = \bar{E}_i - \bar{C}_i / \lambda$$

$$\overline{\text{NMB}}_i = \bar{E}_i \lambda - \bar{C}_i$$

In this chapter we will introduce you to the key mechanisms for summarising the uncertainty in the ICER and Expected Net Benefit from cost effectiveness models, how to construct them and their strengths and weaknesses. In the next section we introduce the scatter plot on the cost effectiveness plane (Black 1990), Sect. 11.3 focuses on the cost effectiveness acceptability curve (CEAC) (Van Hout et al. 1994) and Sect. 11.4 considers the cost effectiveness acceptability frontier (CEAF) (Barton et al. 2008). Section 11.5 takes you through constructing each of these outputs for the outputs for the exercise model. Section 11.6 summarises the material covered in the Chapter.

## 11.2 Scatter Plots on the Cost Effectiveness Plane

Whilst acknowledging the limitations of working in the cost effectiveness plane rather than the Net Benefit Framework, the scatter plot on the cost effectiveness plane is frequently included in published cost effectiveness analyses and in submissions to reimbursement authorities. Experience of working with decision makers also shows us that pictures are often more powerful than words or numbers; thus, scatter plots are considered useful by decision makers for providing insight into the magnitude of uncertainty in the outputs of a CEA, and whether the location of the uncertainty in the outputs translates into significant decision uncertainty; i.e. the risk that making the decision given current evidence carries a significant risk of making the wrong decision. The wrong decision in this context is defined as rejecting a technology that is in fact good value or accepting a technology that is not good value.



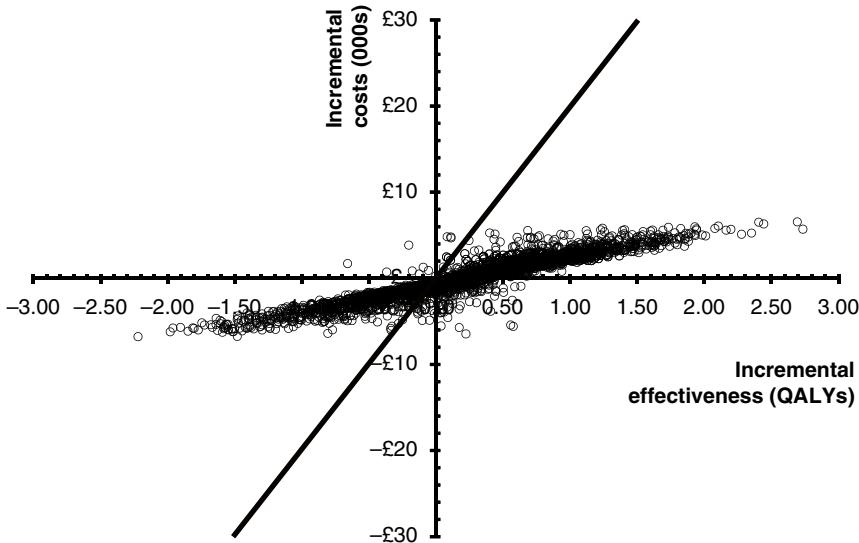
**Fig. 11.1** The cost effectiveness plane: £30,000 per QALY cost effectiveness threshold

To briefly recap the material from Chap. 4, the cost effectiveness plane is defined by the incremental cost on the vertical axis and incremental effectiveness (QALYs) on the horizontal axis (See Fig. 11.1a). There are four quadrants, which are conventionally referred to using compass terminology as the North East, South East, South West and North West. We can locate any possibly output from a CEA in one of these four quadrants.

- In the North East quadrant, we locate ICERs where the new technology produces more health but also costs more than the comparator.
- In the South East quadrant, we locate ICERs where the new technology produces more health and costs less than the comparator.
- In the South West quadrant, we locate ICERs where the new technology produces less health but also costs less than the comparator.
- In the North West quadrant, we locate ICERs where the new technology costs more and produces less health than the comparator.

In the North West and South East quadrants, the decision is obvious because either the comparator or the new treatment is dominated. However, in the North East and South West quadrants the decision makers need to consider their cost effectiveness threshold in order to identify the correct decision. We can plot the cost effectiveness threshold for a given decision maker in the cost effectiveness plane as a line that has the gradient of the cost effectiveness threshold. Figure 11.1 plots a cost effectiveness threshold of \$50,000 per QALY on the cost effectiveness plane. ICERs that fall below the threshold are interpreted as good value and ICERs that lie about the threshold are poor value.

In Chap. 4, we noted that each run of the model produces an estimate of the costs and outcomes from each technology and that it is therefore possible to calculate the



**Fig. 11.2** Scatterplot on the cost effectiveness plane: Oncotype Dx-guided chemotherapy vs. chemotherapy for all in early breast cancer

ICER for each run. The scatter plot on the cost effectiveness plane is the graphical representation of each of run-specific ICERs. In Fig. 11.2, we have plotted the ICERs for 5000 simulations from an analysis of Oncotype Dx-guided chemotherapy compared to chemotherapy for all in the management of early breast cancer. We have also plotted the cost effectiveness threshold where  $\lambda = \text{£}30,000$ .

What does this plot tell us? If we start by considering where the data points are located on the cost effectiveness plane, we can see that the majority of the data points are in either the North East quadrant; i.e. Oncotype Dx is producing more QALYs but at a higher cost, or the South West quadrant. Of the remaining data points, it appears that some suggest that Oncotype Dx may be dominant (cost saving and increased effectiveness, South East). There are also a few data points in the North West quadrant suggesting that Oncotype Dx could, but is unlikely to, both increase costs and reduce health, compared to chemotherapy for all.

On the basis of this diagram alone, it is not possible to tell if Oncotype Dx is likely to be cost effective; the decision maker needs to consider whether the incremental cost is justified by the magnitude of the incremental QALYs across all four quadrants. As a shorthand, we can also consider the location of the data points relative to the line representing points on the cost effectiveness threshold. Visually we can see that many of the data points are above the cost effectiveness threshold; for many of the individual model runs, the results suggest that Oncotype Dx would displace more population health than it would produce and would thus be poor value. Just looking at the scatter plot we cannot say exactly what proportion of the simulations lie above the threshold. As was argued earlier in Sect. 4.8, these sorts of graphs of the cost effectiveness plane are of limited value, and the choice of which option is cost effective comes down to a calculation of net benefit. Here, by balancing expected

costs against expected benefits and the cost effectiveness threshold, an analyst can provide a much clearer illustration of which option is cost effective.

Referring back to Box 11.1, we can see that the Net Monetary Benefit is calculated as the QALYs multiplied by lambda, minus the cost. To compare two or more technologies, we simply compare Net Monetary Benefit (NMB); whichever intervention has the higher positive NMB will be the preferred intervention. The probability that the new technology is cost effective compared to the current technology is calculated by comparing the NMB figures for each simulation run, counting how many times the new technology has the greater NMB and dividing this by the total number of simulations.

### 11.3 Cost Effectiveness Acceptability Curves (CEACs)

In Fig. 11.1 we could not tell how frequently points appeared in each quadrant, but we can obtain some idea of how common they might be in the case where we consider only two options. Presenting only an identified ‘cost effective’ option loses even the limited amount of information provided by investigating a scatter plot on the cost effectiveness plane. As an alternative, the CEAC is a graphical representation of the quantification of the uncertainty around the expected cost effectiveness. It is also the first graphical output from PSA that is understood most easily in the Net Benefit Framework.

In Table 11.1 we have reproduced a subset of ten simulations from the Oncotype Dx analysis to illustrate the calculation of net benefits and CEAC figures. In this data we can see that there is a 70 % chance that Oncotype Dx would be cost effective and a 30 % chance that it would not be, given a cost effectiveness threshold of £30,000 per QALY. To check you understand the calculation, you may wish to reproduce the calculations in NMB columns. The figures in these columns have been rounded up. In reality, decision makers are often uncertain about the actual

**Table 11.1** Uncertainty in the Net Monetary Benefit for Oncotype Dx compared to chemotherapy for all in early breast cancer:  $\lambda = \text{£}30,000$  per QALY

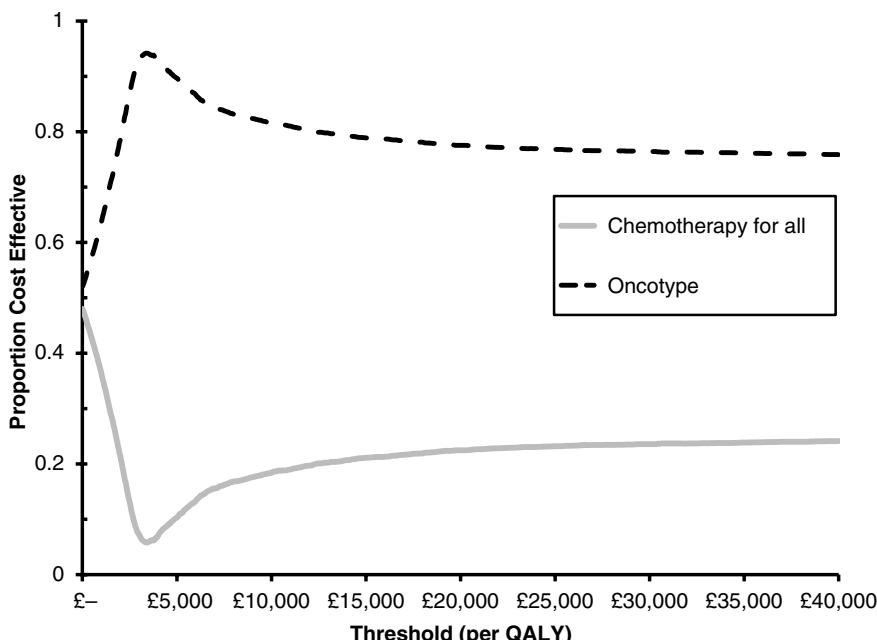
Chemotherapy for all			Oncotype DX		
QALYs	Costs	NMB	QALYs	Costs	NMB
8.664	£14,813	£245,107	8.937	£16,410	<b>£251,689</b>
8.578	£14,587	£242,757	8.729	£14,757	<b>£247,121</b>
8.620	£16,172	<b>£242,428</b>	9.769	£58,432	£234,623
8.497	£19,583	£235,326	8.622	£16,152	<b>£242,520</b>
9.137	£26,555	£247,554	10.516	£51,953	<b>£263,522</b>
8.859	£14,319	£251,444	8.968	£14,511	<b>£254,537</b>
9.335	£17,943	<b>£262,107</b>	10.023	£65,120	£235,570
9.101	£19,845	<b>£253,176</b>	10.325	£57,333	£252,416
9.032	£27,397	£243,569	10.779	£65,485	<b>£257,891</b>
8.828	£19,556	£245,274	9.649	£37,124	<b>£252,352</b>

Highest NMB is **bolded** in each row

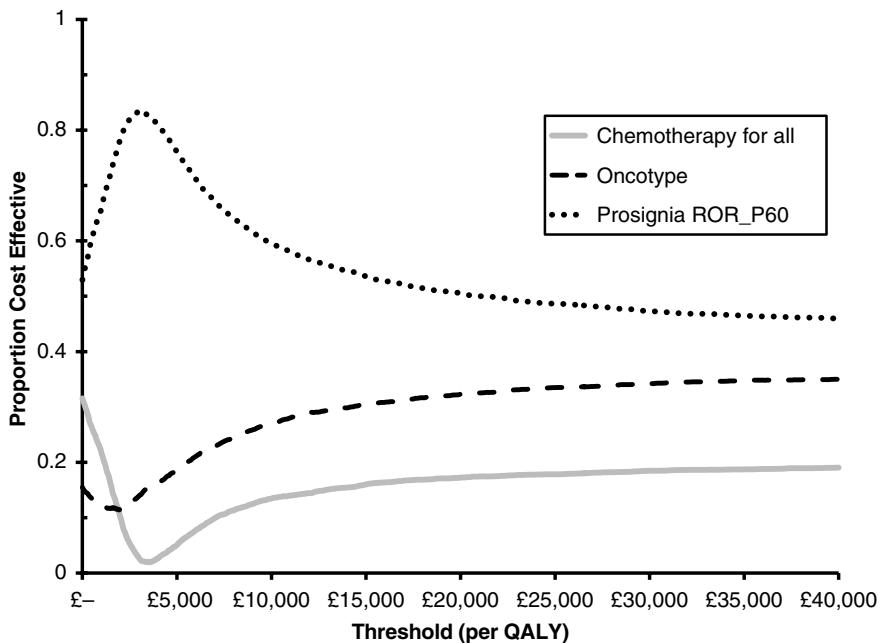
value of the threshold and will therefore be interested in both the Expected NMB and the probability that the NMB is positive over a range of values. In the case of Table 11.1, you can see that Oncotype Dx has the higher NMB in 7 of 10 (70 %) of cases for  $\lambda = £30,000$  per QALY. The CEAC displays this type of information across a range of values for lambda. The CEAC is plotted with probability on the vertical axis and willingness to pay (WTP)/cost effectiveness threshold ( $\lambda$ ) on the horizontal axis. For each value of lambda ( $\lambda$ ), we calculate the proportion of the simulations for which the intervention has the highest NMB.

For our example, Fig. 11.3 plots this information for Oncotype Dx for values of health between £0 and £40,000 per QALY. We can see that over this entire range, Oncotype DX-guided therapy is more likely to be cost effective than providing chemotherapy for all patients. However, even at £30,000 per QALY, there remains about a one in four chance that providing chemotherapy for all would be cost effective. There is clearly substantial residual decision uncertainty about the value of Oncotype Dx. Whilst the scatter plot on the cost effectiveness plane shown in Fig. 11.2 cannot give an accurate indication as to how likely Oncotype Dx is to be cost effective, the CEAC can.

Recall from Sect. 4.8 that the scatter plots on the cost effectiveness plane also struggle to provide any clear guidance where more than two alternatives are considered. The study from which we derived Figs. 11.2 and 11.3 examined a number of chemotherapy-sparing tests to inform the choice of tests to be entered into a large randomised controlled trial of test-guided chemotherapy in early breast cancer. In order to show how the CEAC deals with more than two alternatives, we will



**Fig. 11.3** Cost effectiveness acceptability curve: Oncotype DX-guided chemotherapy compared to chemotherapy for all in early breast cancer



**Fig. 11.4** Cost effectiveness acceptability curves for multiple technologies: chemotherapy for all, Oncotype Dx and Prosignia

consider Prosignia ROR\_P60-guided therapy in addition to the other two options you have already seen.

In Fig. 11.4 we have plotted the CEACs for Oncotype Dx, Prosignia and chemotherapy for all. This is to show that the CEAC does not suffer from the same interpretation problems as the scatter plot on the cost effectiveness plane. We can see that the probability that Prosignia is cost effective when lambda approximately £30,000 is around 47 %, as compared to 34 % for Oncotype DX-guided therapy and around 18 % for chemotherapy for all. CEACs allow us to clearly differentiate the decision uncertainty regarding each of the comparator technologies and how it varies over a range of values for health.

## 11.4 Cost Effectiveness Acceptability Frontiers (CEAFs)

It is important to remember that CEACs do not tell the decision maker which technology has the highest Expected Net Benefit, and this is the criterion that we expect decision makers will use when selecting a cost effective option. The CEAC plots the probability that a technology will have the highest net benefit at a given value of lambda. As the Expected NMB is affected by the magnitude of any differences in net benefit between options, it is not necessarily the case that the technology with highest Expected Net Benefit at that value will also have the highest probability of being cost effective.

The choice facing decision makers is to maximise Expected Net Benefit for a relevant (but possibly unknown) value of lambda. The CEAF is constructed by first plotting the CEACs for the technologies being compared. We then identify the technology with the highest Expected Net Benefit at each value of lambda and for each CEAC trace the portion where the technology also has the highest Expected Net Benefit.

In Fig. 11.5 we have plotted the CEAF for the three chemotherapy regimes: chemotherapy for all, Oncotype DX-guided therapy and Prosignia-guided therapy, from the OPTIMA-prelim study. Note that at around £25 k per QALY, Oncotype DX-guided therapy becomes the most cost effective option despite having a consistently lower probability of being cost effective than Prosignia-guided therapy.

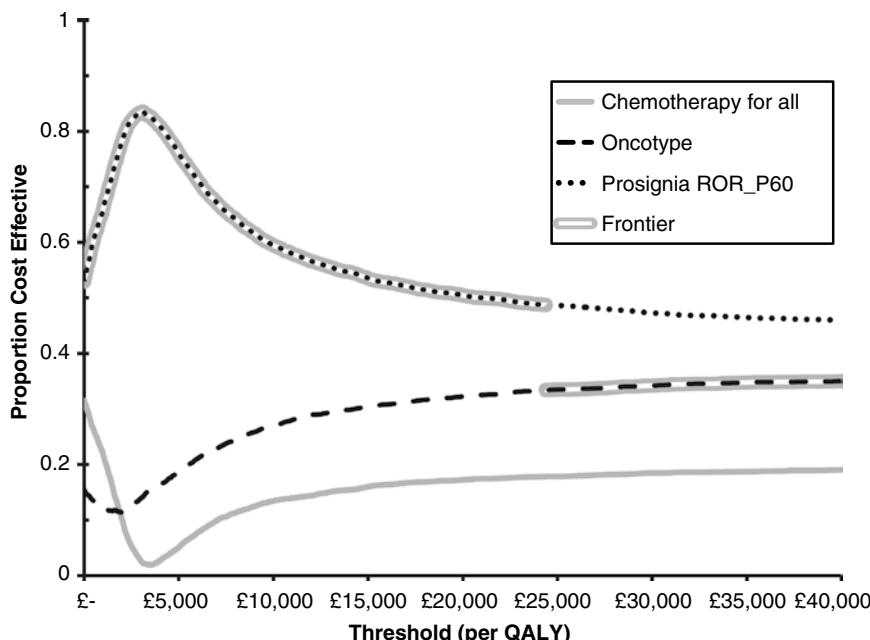
What can we say about the point where the frontier ‘switches’ from indicating that one option is cost effective to indicating that another option is cost effective? At this point, the net monetary benefit of two options is the same. If we call the particular value of the threshold at this point  $\lambda^*$  (the asterisk just means that it is a special value of lambda) then using Box 11.1 we can also say at  $\lambda^*$ :

$$\overline{\text{NMB}}_2 = \overline{\text{NMB}}_1$$

$$\bar{E}_2 \lambda^* - \bar{C}_2 = \bar{E}_1 \lambda^* - \bar{C}_1$$

$$\bar{E}_2 \lambda^* - \bar{E}_1 \lambda^* = \bar{C}_2 - \bar{C}_1$$

$$\lambda^* = \frac{\bar{C}_2 - \bar{C}_1}{\bar{E}_2 - \bar{E}_1} = \text{ICER}$$



**Fig. 11.5** CEAF: Oncotype DX and Prosignia-guided chemotherapies compared to chemotherapy for all in early breast cancer

So, if we can identify the ICERs between the options we consider, we can identify the places where the frontier potentially switches between one decision option and the other.

## 11.5 Scatter Plots, CEACs and CEAf Exercises

Once your simulation has finished running, you are in a position to generate the full range of outputs that probabilistic cost effectiveness models can produce. The remainder of this exercise will require you to construct the following outputs:

- Scatter plot on the cost effectiveness plane
- Cost effectiveness acceptability curve (CEAC)
- Cost effectiveness acceptability frontier (CEAF)

Note that it will be useful later on if you create a variable that records which simulation each cost and outcome estimate is produced by and store this alongside the simulation output data.

The scatter plot on the cost effectiveness plane requires three columns of data: simulation, increment cost Treatment B and incremental QALY Treatment B.

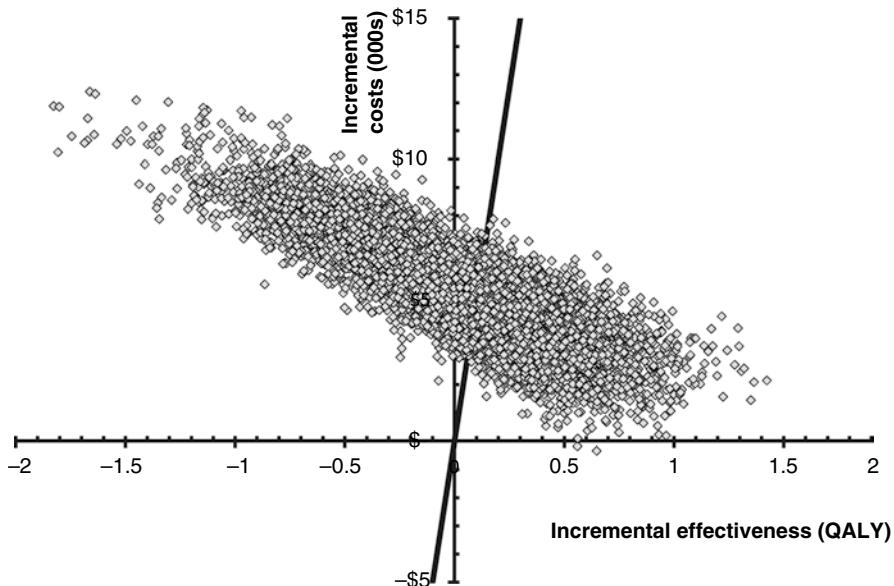
Using the Chart Menu, you need to choose the Marked Scatter as the chart type. Use the dialogue box to insert the incremental cost as the vertical axis and incremental QALY as the horizontal axis. You will then need to give the chart a title and store it in a separate worksheet.

The final task is to add a cost effectiveness threshold line to the chart to aid decision makers with interpreting the results. To give you a hint about how to do this, remember you only need two points to plot a line. The ‘Trendline’ tool might be useful as well. You might want to add the Expected ICER as a third data series onto your scatter plot. We have reproduced our scatter plot below (Fig. 11.6).

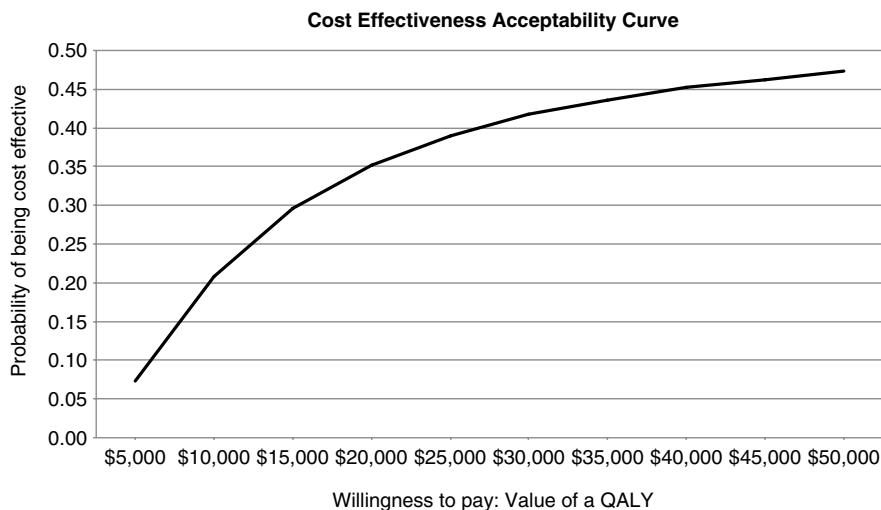
To construct the CEAC, you will need to convert the costs and outcomes for each treatment into Net Benefits and do this for a range of values of lambda. We suggest you use values of lambda between \$0 and \$50,000 in \$5,000 increments; i.e. 11 sets of Net Benefit. Use the line graph function in Excel to plot your CEACs. Remember to label your axes and give the chart a title. Here is our CEAC for Treatment B (Fig. 11.7).

Remember that the CEAF is designed to combine the information contained in the Expected ICER with the insights regarding uncertainty captured by the CEAC. In order to construct the CEAF, you will need to combine the data for the CEACs for both treatments with the data on which has the higher expected Net Benefit at each value of lambda. Normally you will only use a CEAF when you have three or more technologies to compare. Unfortunately, our model only has two comparators; however, the principles for constructing the CEAF work equally well for two technologies. The key requirement is to identify the value of lambda at which the Expected Net Benefit from Treatment B switches to become greater than the Expected Net Benefit from Treatment A.

You will need to construct a variable that records whether Treatment A or Treatment B has the higher Expected Net Benefit. Ideally you will also capture when the Expected Net Benefit is equal. To do this, you will need to define a vari-



**Fig. 11.6** Scatterplot on the cost effectiveness plane



**Fig. 11.7** Cost effectiveness acceptability curve for Treatment B

able using an IF statement. You will need to have this variable for all 10,000 simulation results, assuming you run 10,000 simulations. Once you have done this, you will need to construct another variable that captures the proportion of occasions when Treatment A is preferred to Treatment B. To do this, you will need to use the COUNT and COUNT IF functions.

Having set up the variables for identifying what proportion of the time Treatment B is expected to be preferred to Treatment A, the key information for the CEA, you now need to set up an array that will capture this value for a range of values of lambda and the data necessary to plot the CEACs and CEA. We recommend that you create five new variables – lambda, Treatment\_A\_%; Treatment\_B %, Treatment\_A\_NB and Treatment\_B\_NB. These variables will pick up the results from each lambda-specific simulation. You will need to create an array into which the results can be copied so that they are available for plotting the CEACs and CEA. You may wish to write a macro that generates results for each value of lambda and records them in the array you have created. In our analysis we generated data for values of lambda between \$0 and \$150,000 in \$1000 increments.

Once you have generated the data to plot the CEACs, you need to create your CEA variable. Remember this is going to plot the probability that Treatment A is cost effective for those values of lambda where the expected Net Benefit for Treatment A is greater than Treatment B; and the probability that Treatment B is cost effective when it has the higher expected Net Benefit. You have captured the necessary information in the array, so now you need to define the CEA data series. An IF statement will allow you to choose the CEAC data point you need conditional upon which treatment has the higher Expected Net Benefit.

Having defined the three constituent data series for your CEAC plot, the two CEACs and the CEA, you can now use the Insert Line Graph facility within Excel to create the necessary diagram. Save your model into its own worksheet. Remember to label the axes and give the chart a title. Figure 11.8 is the CEA from our model. Does yours look like this? If not, how does it differ and do you know why?

Whilst this chapter covers the major steps in creating the different outputs from your probabilistic model to provide additional support, we have a series of worked exercises using Excel Workbooks available from <http://medhealth.leeds.ac.uk/costeffectivenessmodelling>. A total of five exercises are available as part of the

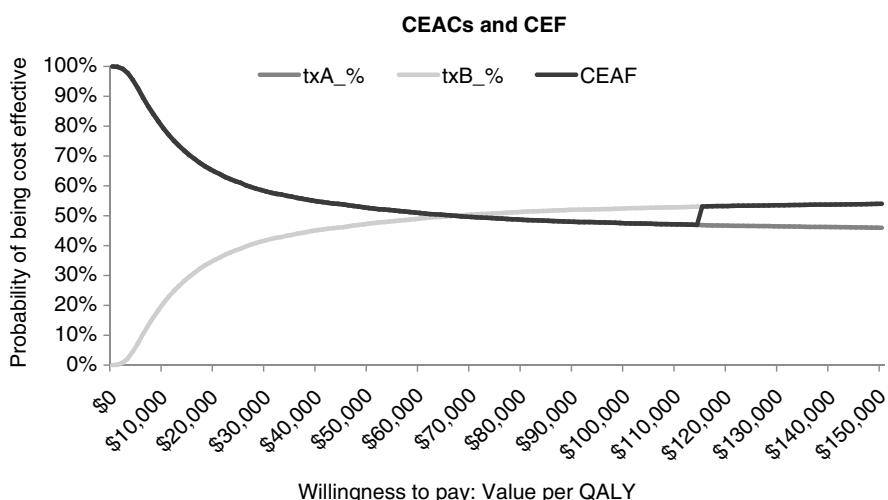


Fig. 11.8 Cost Effectiveness Acceptability Curve

online material. Exercises 11A and 11B use the decision tree based on Evans et al. (1997) that was used in Chaps. 1 and 2, whilst Exercises 11C–11E use the Markov model used in Chaps. 9 and 10.

Exercises 11A and 11B deal with the calculation of the outcomes from a probabilistic sensitivity analysis, with 11A showing the importance of correctly interpreting mean values when dealing with ICERs and 11B giving practice in the calculation of net benefit figures.

Exercise 11C–11E outline how the diagrams like those identified in this chapter can be produced (for a different example), showing the cost effectiveness scatter plot (11C), and cost effectiveness acceptability curves (11D) and cost effectiveness acceptability frontier (11E).

## 11.6 Summary

- The outputs from PSA are used to construct the incremental cost Effectiveness ratio (ICER) using the expected values of the costs and outcomes distributions for each intervention being compared.
- The ICER does not provide decision makers with information on the uncertainty around the expected costs and outcomes.
- Scatter plots on the cost effectiveness plane graph the ICER for all model simulations, providing a visual account of the absolute and decision uncertainty in the ICER.
- Scatter plots on the cost effectiveness plane are not quantitative and therefore are often problematic to interpret, especially when there are more than two interventions being compared.
- The Net Benefit Framework eliminates the problem of interpretation of ICER results and allows the quantification of the decision uncertainty for a given cost effectiveness threshold ( $\lambda$ ).
- The CEAC plots the probability that a technology is cost effective for a range of values of  $\lambda$ .
- The CEAF combines the CEAC with a plot of the technology with the highest Expected Net Benefit over the same range of values of  $\lambda$ .
- In our Exercise model, we constructed the scatter plot on the cost effectiveness plane, the CEAC and the CEAF, using the outputs of the simulations we created in Chap. 10.

## References

- Barton GRB, Fenwick AH, EAL (2008) Optimal cost effectiveness decisions: the role of cost effectiveness acceptability curve (CEAC), the cost effectiveness acceptability frontier (CEAF) and the expected value of perfect information (EVPI). *Value Health* 11(5):886–897

- Black WC (1990) The CE plane: a graphic representation of cost-effectiveness. *Med Decis Making* 10:212–214
- Evans KW, Boan JA, Evans JL, Shuaib A (1997) Economic evaluation of oral sumatriptan compared with oral caffeine/ergotamine for migraine. *PharmacoEconomics* 12(5):565–577
- Stinnett AA, Mullahy J (1998) Net health benefits: a new framework for the analysis of uncertainty in cost effectiveness analysis. *Med Decis Making* 18(Suppl):S68–S80
- Van H, Al BA, Gordon MA, Rutten GS, FFH (1994) Costs, effects and C/E ratios alongside a clinical trial. *Health Econ* 3(5):309–319

# Chapter 12

## Investing in Health Care, Research and the Value of Information

**Abstract** The evidence base for most technologies are highly uncertain at the time of regulatory approval. One response to this uncertainty is to regulate access to technologies using risk-sharing and patient access schemes. This chapter describes how net benefit probability maps can help decision makers understand the impact of access with evidence development schemes on both value and decision uncertainty. The chapter also illustrates how value of information analyses can allow decision makers to compare different responses to uncertainty, with examples that compare the value of delaying reimbursement to allow future research with the use of different regulatory schemes that may allow at least limited access to technologies until research reports.

### 12.1 Introduction

At this point, you know how to produce the standard range of outputs from a probabilistic model: the expected incremental cost effectiveness ratio (ICER), the Expected Net Benefit, scatter plot on the cost effectiveness plane, cost effectiveness acceptability curve (CEAC) and cost effectiveness acceptability frontier (CEAF).

In Chap. 5, we considered the need to discount costs and benefits to capture individuals' time preference, i.e. the reality that a \$100 today will be preferred to \$100 next year, and avoiding an ill-health event in the future will not be valued as highly as avoiding an ill-health event today. Once we have applied the appropriate discount rate, we report costs and benefits in *present value* terms. The outputs of cost effectiveness analyses considered so far do not provide decision makers any information on when the costs are incurred nor when the benefits are received. In this chapter, we return to the question of the temporal distribution of costs, benefits and uncertainty over time, how this information can still be relevant to decision makers beyond the question of time preference and how it is possible to present this information in an accessible manner for decision makers. Specifically, we will introduce three additional outputs from probabilistic sensitivity analysis (PSA): the net benefit breakeven curve, the net benefit probability map (NBPM) (McCabe et al. 2013) and the value of perfect information (Claxton et al. 2002). The first two allow the decision maker to see how costs and benefits of new technologies are distributed over time and the differences in the

magnitude and location over time of the uncertainty in the expected costs and outcomes of the interventions. The third output, the value of information (VOI), supports decision makers' understanding of the value of delaying a reimbursement decision to allow more evidence to be generated. The final exercise at the end of this chapter requires you to produce all three outputs for the model that you have constructed. Section 12.2 reviews the increasingly sophisticated approaches that decision makers are adopting to the evaluation and reimbursement of new health technologies. Section 12.3 introduces the idea of considering reimbursement decisions as an investment and looking at how long we should expect it to take for a technology to break even in the Net Benefit (NB) Framework. Section 12.4 extends this to take account of the uncertainty in the time to break even using the NBPM, exploring how the effect of different patient access schemes can be shown to decision makers using the NBPM. Section 12.5 introduces the concept of the VOI and how to calculate the expected value of perfect information (EVPI). Section 12.6 then considers the disaggregation of the VOI using the expected value of perfect parameter information (EVSSI) and the expected value of sample information (EVSI). Section 12.7 contains the final exercises, requiring you to construct the NBPM and the EVPI for the model. Section 12.8 provides a summary of the material covered in the chapter.

## 12.2 Uncertainty and Health-Care Reimbursement Decision-Making Processes

High levels of uncertainty are inherent in the evidence base for innovation technologies at the time of regulatory approval. The number of patients who have been exposed to the innovation is small as a proportion of the population who will likely receive it if it is funded. The duration of the follow-up for patients on therapy will typically be very short compared to the length of time that patients will live with therapy once it is implemented. Understanding of the impact of casemix and health-care setting on the effectiveness and safety of the technology is an important determinant of the real-world value of innovations; however, for good reasons, the premarket evidence development process typically includes a highly select patient group and takes place in similarly select health-care environments. In contrast, the comparator technology in a cost effectiveness analysis is likely to have been in routine clinical use for a considerable length of time and subject to a much higher volume of both efficacy and pragmatic effectiveness research. As a result, the uncertainty in the estimated NB is unequally distributed between the innovative technology and current standard of care; and whilst the technology that will be displaced elsewhere in the health-care system through opportunity cost is unknown, as a currently used technology, the same insights suggest that there will be greater certainty about its effectiveness and safety, even when we don't know what the displaced technology will be. As a result, the decision uncertainty for adopting innovative technologies relates primarily to the innovative technology.

Reimbursement authorities in most, if not all, developed health-care systems have responded to the challenge of decision uncertainty for innovative technologies through various forms of risk-sharing and patient access schemes. Stafinski and

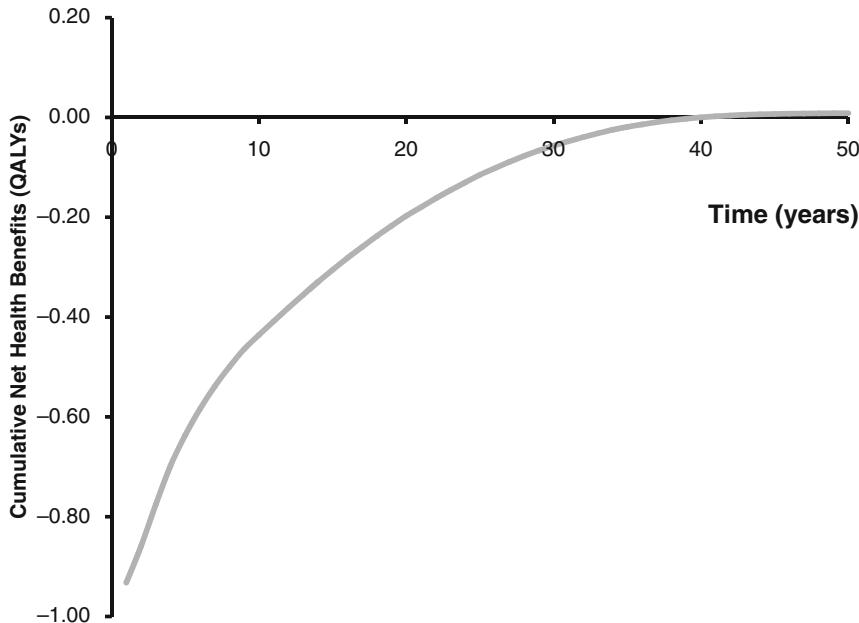
colleagues' (2010) review of these schemes reported that although their use was widespread, the evidence that they protected health system budgets from inefficient use or that they generated new evidence on the real-world value of the technologies in a timely manner was extremely limited. In part, the failure of these schemes to meaningfully address the decision uncertainty is explained by the fact that all risk associated with the innovative technology is transferred to the payer when the technology is paid for, which almost invariably takes place at the time of its first administration. Whilst the total decision uncertainty may be reduced by a reduction in the expected price, the remaining risk is not shared between the payer and the manufacturer of the technology. The model outputs we have described so far – ICER, NB, CEAC and CEAF – do not communicate this information to the decision maker. At best, they describe the change in aggregate decision uncertainty with and without the 'risk-sharing' scheme. More nuanced presentations of the location of decision uncertainty are required if analysts are to support decision makers in developing more informed reimbursement strategies.

## 12.3 Investing in Innovative Health Technologies

In many areas of public and private life, we are faced with opportunities that involve current sacrifices to fund activities with the expectation that over time the activity will produce greater things that we value more highly than whatever was initially sacrificed; because the expected benefits are in the future, there is some degree of uncertainty around the expectation. We refer to such opportunities as investments.

Many years ago, Kenneth Arrow observed that uncertainty is a key characteristic of health care as a commodity, specifically noting the uncertainty that an individual patient will benefit from a treatment (Arrow 1963). It is surprising therefore that cost effectiveness analysis and health technology assessment more broadly have made such limited use of the analytical toolkit applied to other investments. One of the standard tools for exploring the distribution of investments (costs) and returns (benefits) over time is the breakeven curve. Conventional investments characterise the breakeven point in terms of money. However, health-care budgets are not invested to create money; they are invested to create health. Hence, the appropriate metric or index for exploring the distribution of investment and returns over time for a health system is the Net Health Benefit (NHB). Figure 12.1 plots an illustrative net benefit breakeven curve (NB\_BEC). The NB\_BEC is obtained by plotting the cumulative Expected NB at the end of each model cycle over the time horizon of the model.

Conventional investment breakeven analysis uses expected values to identify the time when the value of the returns is expected to exceed the value of the initial investment. Uncertainty is largely addressed through the construction of a portfolio of investments with a range of individual risk profiles. Health technology reimbursement decisions are made as a series of single investments at the margin of the health-care budget, and the decision maker has little, if any, control over which investments opportunities are presented for consideration. A portfolio approach to managing uncertainty is not available; hence, the investment-specific uncertainty is key to supporting the decision maker in managing risk.



**Fig. 12.1** Net benefit breakeven curve

**Table 12.1** Outputs from simulation model – cycles 1–5

Cycle	Treatment A (\$s)	Treatment A QALYs	Treatment B (\$s)	Treatment B QALYs
1	\$16,878	0.405	\$30,665	0.596
2	\$17,893	0.404	\$30,424	0.640
3	\$16,931	0.532	\$28,499	0.699
4	\$15,646	1.307	\$29,998	1.524
5	\$17,258	-0.609	\$31,527	0.120

Stochastic cost effectiveness models generate the data to provide a comprehensive description of the uncertainty at each time point over the model time horizon. However, for each simulation, standard practice is to capture only the expected value of the costs and outcomes over the time horizon of the model. The cycle-specific costs and outcomes for each intervention are generated but not captured. Without capturing these data for each cycle, it is not possible to describe how the uncertainty in the time to break even for the health technology investment.

Consider the Markov model you have constructed. The simulation process draws a sample value from each parameter distribution and then uses the drawn values to simulate the progression of a cohort of two patients through the model: one cohort receiving standard care and the other cohort receiving the new (innovative) technology. Table 12.1 shows an edited version of the output table for cycles one through 5 for one simulation. It is clear from this table that for each cycle, we have generated the data necessary to calculate a cycle-specific NB for each technology and hence

**Table 12.2** Outputs from simulation model – cycles 1–5

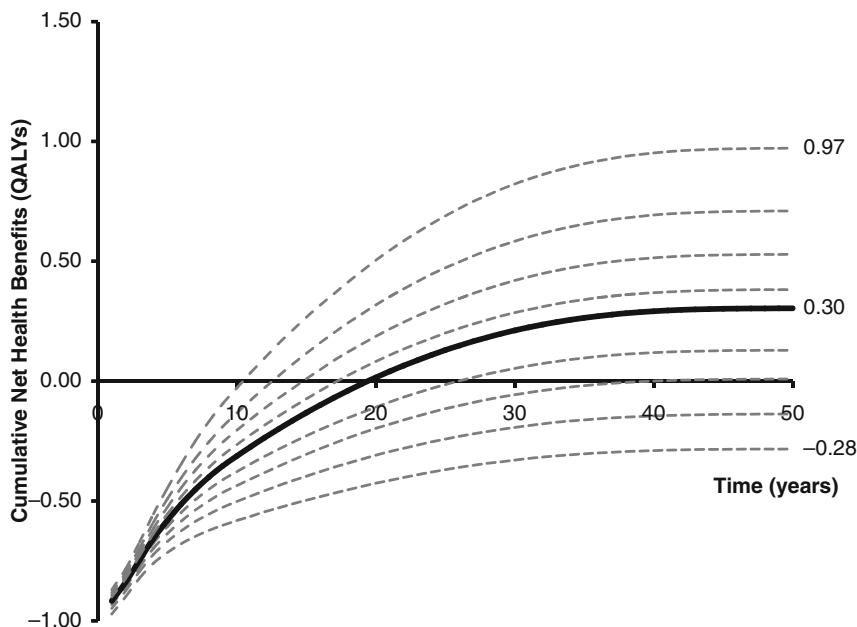
	txA_cost	txA_QALY	txB_cost	txB_QALY	Incremental QALYs	Incremental costs	INB	Cum INB
1	\$16,878	0.405	\$30,665	0.596	0.191	\$13,788	-0.085	-0.085
2	\$17,893	0.404	\$30,424	0.640	0.236	\$12,531	-0.015	-0.099
3	\$16,931	0.532	\$28,499	0.699	0.167	\$11,569	-0.064	-0.163
4	\$15,646	1.307	\$29,998	1.524	0.216	\$14,351	-0.071	-0.234
5	\$17,258	-0.609	\$31,527	0.120	0.730	\$14,269	0.444	0.210

the Incremental Net Benefit (INB) for Treatment A compared to Treatment B, assuming we know the value of lambda, as well as a cumulative NB overall the model cycles. Table 12.2 shows these additional outputs assuming lambda is \$50,000 per QALY.

The data presented in Table 12.1 is for a single draw of parameter values for all parameters in the model. When we repeat the process with a new draw, we will get different values for all of the cycle-specific outputs. When we repeat the draw process thousands of times, we are able to produce a probability distribution for each of these cycle-specific outputs. Hence, we can quantify the uncertainty in the Expected NB for each technology and the INB at every time point over the time horizon of the analysis, at a granularity determined by the cycle length. Having said this, the quantity of data generated is extremely large. Consider a model with 50 cycles and a probabilistic sensitivity analysis running 5000 simulations; four outcomes (costs and outcomes for each of the two technologies) will generate 1 million items of data, which quantify the expectation and uncertainty of the cumulative INB over time. In order to make the information accessible to decision makers, we use NBPM (McCabe et al. 2013).

## 12.4 Net Benefit Probability Map and Managing Decision Uncertainty

The NBPM uses the idea of contours to represent the uncertainty around the cumulative NB and uncertainty in the time to break even. For each time point, it is possible to construct a probability distribution for the Cumulative Incremental NB, from the PSA data. In this distribution at each time point, we can identify any particular centile value. McCabe et al. (2013) suggest using deciles. Joining the deciles at each time point creates uncertainty contours in the Incremental Net Benefit – Time Plane. For a model with 50 cycles, this requires 450 data points in the Incremental Net Benefit – Time Plane – a considerably more tractable data burden than the millions of data points created by the full simulation. By plotting the expected cumulative Expected NB (the NB\_BEC) on the same plane as the contours, the analyst can show decision makers how uncertainty is distributed over time, the uncertainty in the time to break even and the risk that investing in the new

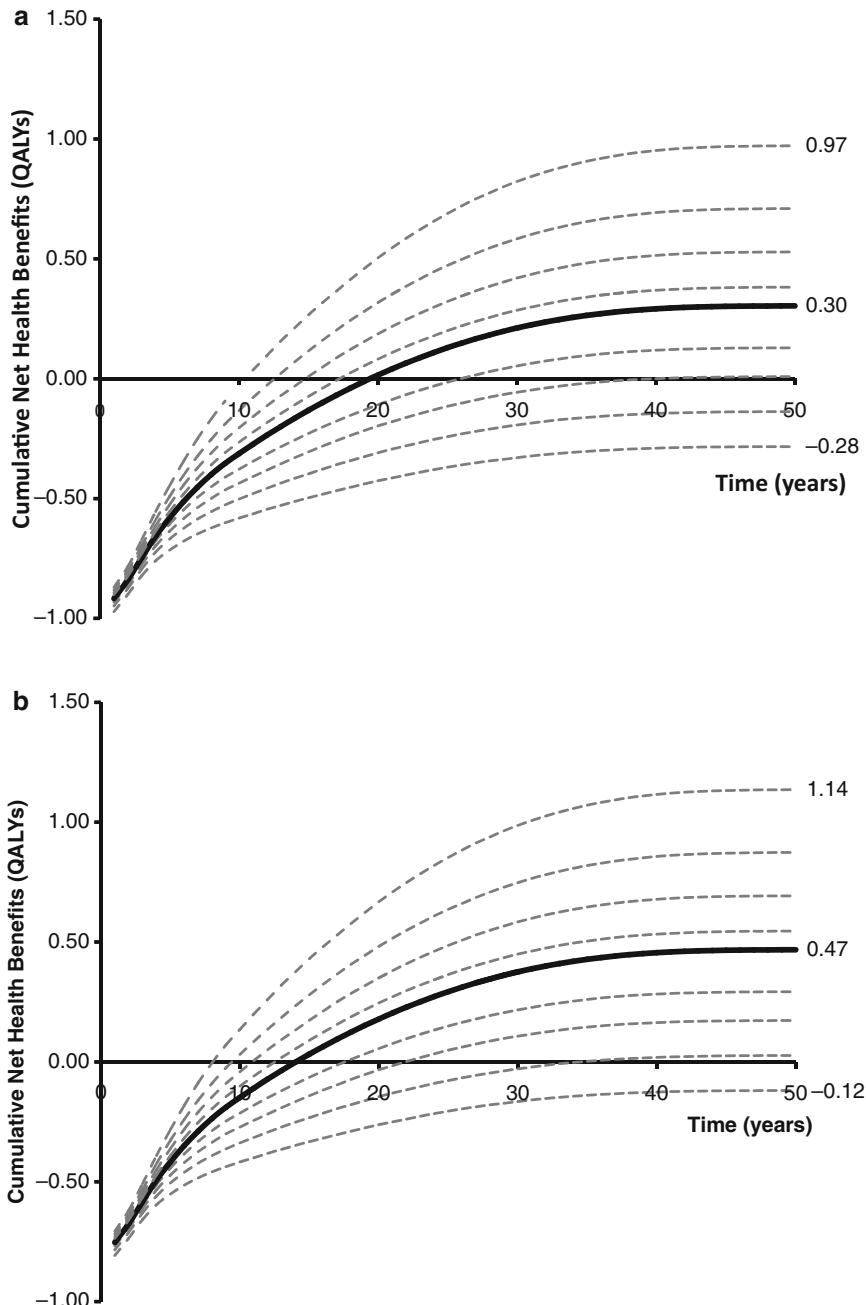


**Fig. 12.2** A net benefit probability map

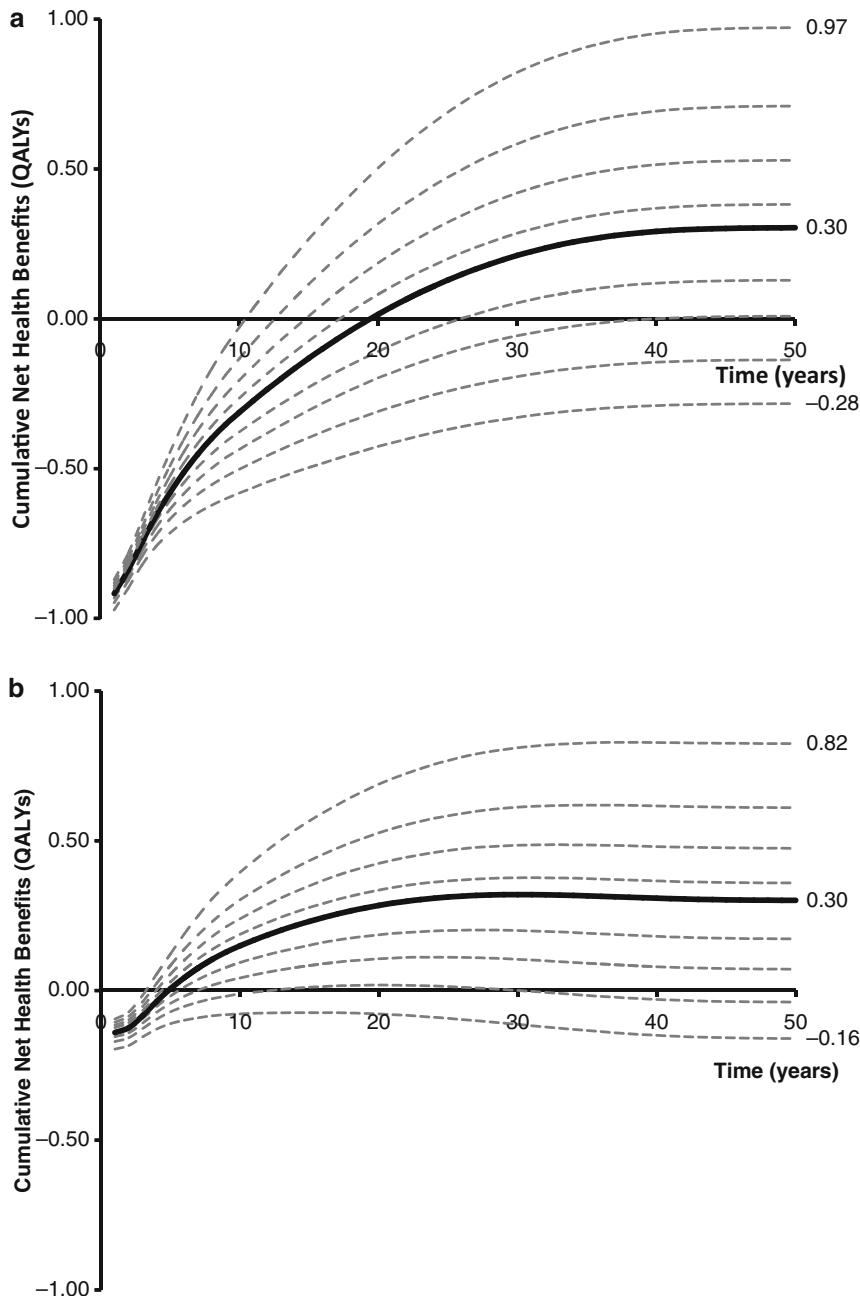
technology will never break even. Figure 12.2 shows a NBPM with the deciles of the Cumulative NB distribution at each time point highlighted.

Earlier in this chapter, we outlined how the vast majority of schemes designed to reduce the risk associated with investing in innovative technologies have been less effective than hoped. This is primarily because they have tended to consist of a package of price reductions and evidence development to inform future decisions. In the vast majority of cases, the data collected has not impacted upon the price paid for an individual patient's use of the technology. The agreed price has been paid in full at the time of administration, and all the risk associated with the uncertainty around the value of the technology has passed to the health-care payer at that point. Figure 12.3 shows how a price reduction impacts upon the distribution of uncertainty and the expected time to break even. We can see that the scale of the initial investment is less, the expected time to break even is reduced and the risk that the investment will never break even is reduced. However, the pattern in the uncertainty is little changed between the original price and the discounted price. Hence, a price reduction, such as that implemented in the UK Multiple Sclerosis Risk-Sharing Scheme (UK Department of Health 2002), does not benefit the health payer by reducing uncertainty, but by reducing the scale of the upfront investment.

Figure 12.4 shows the impact of linking payment to performance over time, at the individual patient level. It can be seen that this approach radically changes the distribution and magnitude of uncertainty. The investment is expected to break even earlier, and the risk that it will not break even is also reduced (Edlin et al. 2014).



**Fig. 12.3** Impact of price discounts on time to break even and uncertainty



**Fig. 12.4** Impact of Pay for Performance on time to break even and uncertainty

The NBPM will support decision makers in assessing alternative patient access schemes in a much more nuanced manner, showing the impact on the Expected NB, the uncertainty around that expectation, the time to break even and the uncertainty around the expected time to break even. By contrast, the most frequently used approach to communicating uncertainty in cost effectiveness analyses, the CEAC, would only show the impact on the uncertainty around the Expected NB at the end of the model time horizon.

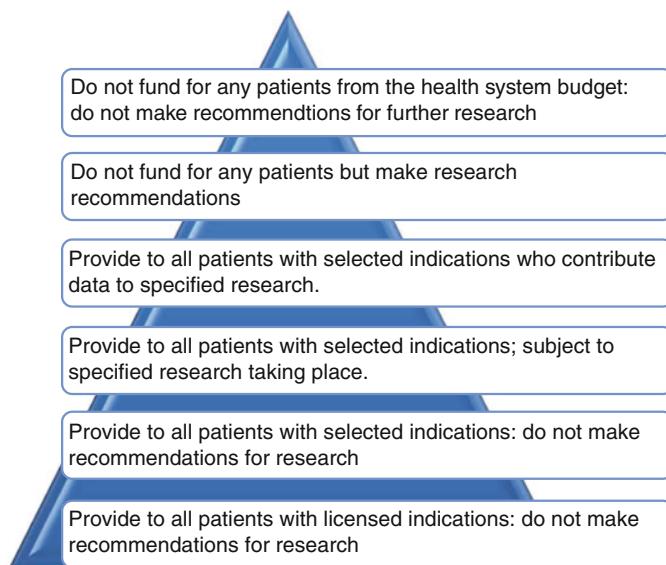
## 12.5 Delaying a Reimbursement Decision for More Research

Historically, health-care reimbursement decisions have been dichotomous – yes or no – with management of budget impact of positive decisions via the identification of patient subgroups within the total population who could clinically benefit, for whom funding would be available. However, in the last decade, there has been a significant expansion in the range of decision options available to decision makers. Figure 12.5 illustrates the reimbursement decision options as a pyramid, with full reimbursement as per clinical indications as the foundation. As decision makers move up the pyramid, reimbursement becomes more conditional, and the data collection requirement becomes more substantial. There are two distinct types of patient access schemes: frequently referred to as Only with Research (OWR) and Only in Research (OIR) (McCabe et al. 2010). OWR schemes reimburse a technology subject to clearly specified data to be collected; however, individual patients' access is not dependent upon them providing data for the research study. In contrast, OIR studies require that patients who receive the technology generate data for the associated research study.

There are two further decision options – rejection of reimbursement but with recommendations for research that would be expected to impact upon a future review of the reimbursement decision but without a commitment to pay for the technology in the context of such research. Implicit in this is an assessment that costs of implementing the recommended research would not be a good investment from the perspective of the health system. The final option is an absolute ‘no’ without any research recommendations. Implicit in this decision option is a judgement that there is no further research that could realistically change the decision.

Originally, the move toward conditional reimbursement was developed through ad hoc responses to difficult reimbursement decisions, characterised by substantial clinical need but high-cost technologies imposing large budget impact on health system budgets and highly uncertain evidence on which to base the investment decision – see, for example, the UK Multiple Sclerosis Risk-Sharing Scheme (UK Department Health 2002). As these schemes – often called risk sharing or access with evidence development (AED) schemes – became popular with health systems (Stafinski et al. 2010), researchers began to investigate whether they had achieved their stated aims and whether there was scope for more formal approaches to their design (Walker et al. 2012).

Stafinski and colleagues undertook an extensive review of the published and grey literature, identifying a substantial number of AED schemes, and concluded that



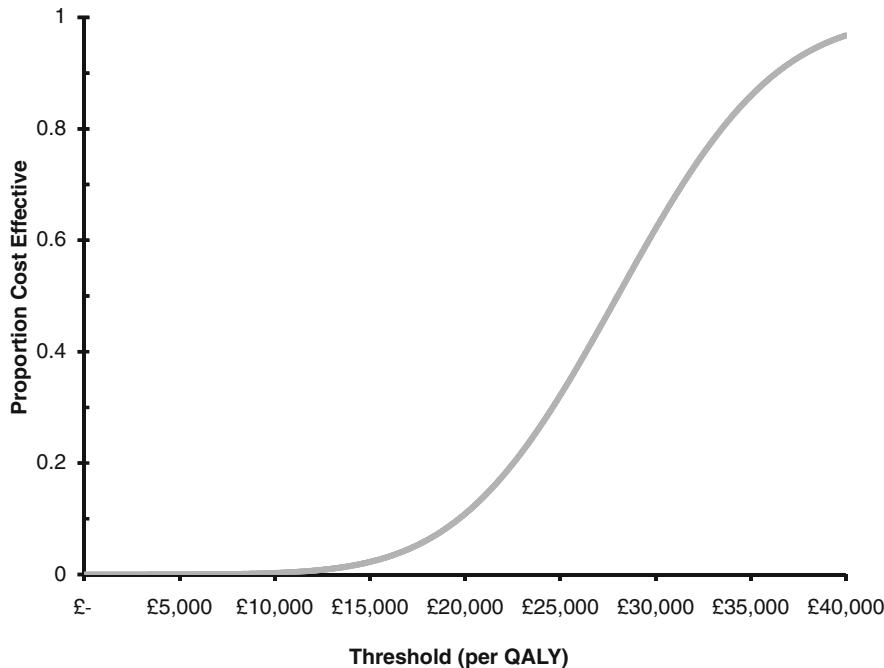
**Fig. 12.5** Reimbursement decision options (Based upon Edlin et al. 2014)

few, if any, had achieved their stated objectives of providing access for patients whilst improving the evidence base for future decisions. The vast majority of these schemes fall into the OIR category.

Other authors considered whether it would be possible to take a prospective approach to the design of AEDs, by linking the cost of developing additional evidence through research to the value of that evidence through reducing the risk of making the wrong decision (e.g. Walker et al. 2012). The risk of making the wrong decision is called decision uncertainty, and the methods for quantifying decision uncertainty and evaluating research according to their impact on decision uncertainty are known as value of information (VOI) analyses. The remainder of this chapter provides an overview of the foundation concepts of VOI (Claxton et al. 2001) and the implementation of these concepts using stochastic decision analytic cost effectiveness models.

### ***12.5.1 Uncertainty in Decision Making and the Cost of Making the Wrong Decision***

The simplest way to grasp the concept of decision uncertainty is to examine the CEAC from an analysis of two alternative courses of action (Fig. 12.6). If the willingness to pay (WTP) for an additional quality adjusted life year (QALY) is £25,000, then we can read off that the probability that this is cost effective is 70%. The dual of this is that there is a 30% chance that that this technology would actually displace more population health than it would produce. The decision uncertainty around this technology is therefore the 30%. The expected costs (per person treated) of making the wrong decision by reimbursing this technology is calculated as  $0.3 * £25,000 = £7,500$ .



**Fig. 12.6** Cost effectiveness acceptability curve and decision uncertainty

Assuming that 10,000 people would receive the technology, the expected total population cost of making the wrong decision would be £75 million.

### 12.5.2 *Expected Value of Perfect Information and the Value of Sample Information*

The concept of Expected VOI starts with the observation that the reimbursement decisions are made prospectively and that outcomes are captured retrospectively. Thus, decisions are always made with the risk that the expected outcome will not be observed in practice. There is no way around this, and therefore, decisions have to be made on the basis of the decision makers' expectation of what will happen given the best available evidence.

Table 12.3 illustrates this idea. Each row represents a realised state of the world following a positive reimbursement decision for two treatments: Treatment A and Treatment B. If the decision maker has to choose between A and B, they will use the expectation – in this case, the Expected NHB for Treatment A is 12 and compared with 13 for Treatment B. Therefore, the rational decision maker will choose to fund Treatment B.

Consider a mythical decision maker who actually knows how the world would turn out at the time of making their decision. Such a decision maker would always make

**Table 12.3** Reimbursement decisions under conditions of uncertainty

	Treatment health benefit	
	A	B
State of the World 1	9	12
State of the World 2	12	10
State of the World 3	14	20
State of the World 4	11	10
State of the World 5	14	13
<i>Expectation</i>	12	13

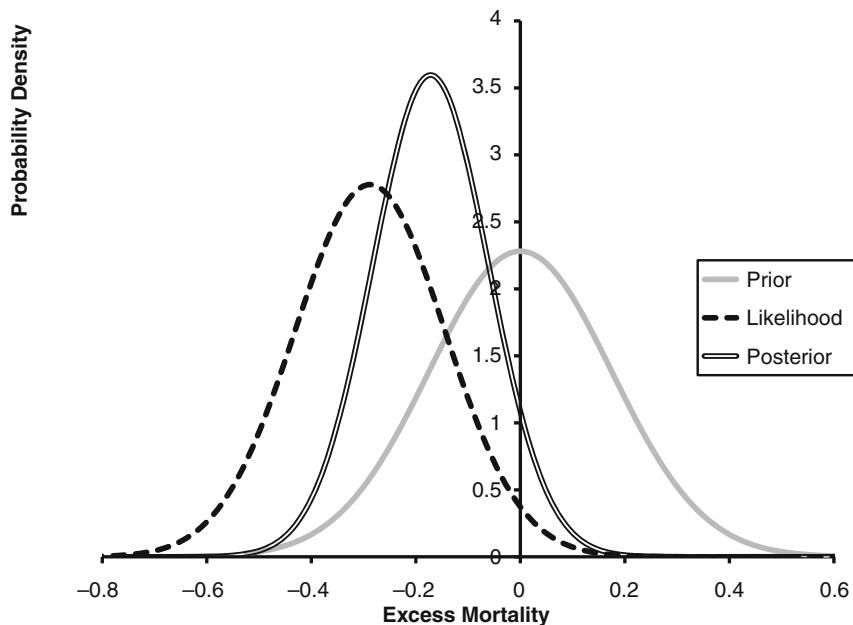
**Table 12.4** Reimbursement decisions with perfect knowledge

	Treatment health benefit		Optimal choice	Maximum health benefit	Health loss
	A	B			
State of the World 1	9	12	B	12	0
State of the World 2	12	10	A	12	2
State of the World 3	14	20	B	20	0
State of the World 4	11	10	A	11	1
State of the World 5	14	13	A	14	1
<i>Expectation</i>	12	13		13.8	0.8

the correct decision. They would always choose the treatment that provided the greatest NB. Hence, in State of the World 1, this decision maker would choose Treatment B, but in State of the World 2, they would choose Treatment A. They would also choose Treatment A in State of the World 5, but in the remaining States of the World, they would choose Treatment B. These choices are illustrated in Table 12.4, and we can see that the Expected NHB for the perfectly informed decision maker is 13.8. The perfectly informed decision maker has an Expected NHB that is 0.8 higher than the real-world decision maker who uses the a priori expectation, 13.8 minus 13. If the function of research is to reduce decision uncertainty, then the value of further research cannot exceed the value of certainty. In this example, the maximum expected value of perfect information (EVPI) is represented by the additional NHB of 0.8. If the value of health is £20,000, as in the previous example, this equates to £16,000 per patient treated. Assuming that it is the same 10,000 patients as in the previous example, the expected population value of perfect information is £160 million.

Obviously, we know that the perfectly informed decision maker does not exist and that the decision maker can never be certain because decisions are about future events and the future however likely can never be certain. However, attaching a value to the perfect information provides a reference point for thinking about the value of additional information. Conceptually, the value of any additional information can be thought of as the reduction in the value of perfect information that it produces.

We can illustrate this by extending this same example. If you look again at Table 12.4, you will see that the majority of the value of perfect information is driven by the possibility that State of the World 2 is realised. States of the World 1, 3 and 4 all support the choice of Treatment B, and thus, further information on the advantage of Treatment B over Treatment A has no value because it would not



**Fig. 12.7** Tri-plot: excess mortality of Treatment B compared to Treatment A

change the decision from that which is made on the basis of the expectation. State of the World 2, should it be realised, leads to a loss of 2 units on the NHB Scale, whilst State of the World 5 would lead to a loss of 1 unit.

For the sake of illustration, we are going to assume that this reflects a small but meaningful possibility that Treatment A has a lower mortality rate than Treatment B. Figure 12.7 is a Tri-plot, showing three distributions on the hazard ratio for mortality on Treatment A compared to Treatment B. The blue distribution reflects our beliefs given the currently available evidence – this is called our Prior. This is the effectiveness distribution in the stochastic analyses from which States of the World 1–5 were sampled.

Concerned about the risk of making the wrong decision, the health system commissions a head-to-head trial of Treatments A and B. The distribution for the relative risk of mortality observed in the trial is described by the light grey curve – this is called the likelihood distribution. When the trial data is synthesised with the prior knowledge (the black curve), the best estimate of the relative risk is captured by the mid-grey curve (the Posterior distribution). This updated distribution is called the Posterior. The important thing to note is that the probability that Treatment B is associated with an excess mortality, according to our Posterior, is much lower – a much smaller proportion of the Posterior distribution is positive compared to the Prior distribution.

Table 12.5 illustrates how the use of the Posterior distribution does our estimation of the EVPI. In State of the World 2, Treatment A has a NHB of 10, whilst Treatment B now has a NHB of 11. Treatment B is now the correct choice in State

**Table 12.5** Impact of additional information on expected value of perfect information

	Treatment health benefit		Optimal choice	Maximum health benefit	Health loss
	A	B			
State of the World 1	9	12	B	12	0
State of the World 2	10	11	B	11	0
State of the World 3	14	20	B	20	0
State of the World 4	11	10	A	11	1
State of the World 5	14	13	A	14	1
Expectation	11.6	13.2		13.6	0.4

of the World 2, and there is thus no Health Loss in this State of the World attributable to using the Expected NHB to choose between Treatments A and B. The expected net present value of perfect information is now only 0.4 on the NHB Scale. Research has reduced the value of perfect information and that reduction gives us insight into the value of the research. Note that the formal calculation of the value of a specific piece of research – the EVSI – is actually more complex than this conceptual illustration. We work through it later in this chapter.

### 12.5.3 Calculating the Expected Value of Perfect Information

Calculating the EVPI is as straightforward as the illustration in Table 12.2. However, it is useful to start with the formal algebra.

Equation 12.1 characterises the Expected NB from treatment (j) given current information theta ( $\theta$ ).

$$\max_j E_\theta NB(j, \theta) \quad 12.1$$

The Expected NB given current information is obtained by choosing the option with the higher Expected NB across all the simulations, j in set theta.

Equation 12.2 characterises the Expected NB from treatment (j) given perfect information.

$$E_\theta \max_j NB(j, \theta) \quad 12.2$$

The Expected NB, given Perfect Information, is obtained by choosing the intervention with higher NB in each simulation and calculating the Expected NB over all simulations in the set theta, that is, the mean value of the higher net benefit from each simulation. This calculation captures the Expected NB if, in every case, we chose the better technology.

The EVPI is simply the difference between the Expected NB given perfect information and the Expected NB given current information.

$$EVPI = E_\theta \max_j NB(j, \theta) - \max_j E_\theta NB(j, \theta) \quad 12.3$$

## 12.6 Disaggregating the Value of Information: Expected Value of Perfect Parameter Information and the Expected Value of Sample Information

Decision uncertainty and the EVPI consider the value of the uncertainty in all the parameters in the decision model. In reality, some parameters are much more uncertain than others. For example, costs of physical therapies in a health-care system are often known with more confidence than the effectiveness of a new drug. This is because the system will typically provide hundreds of thousands of physiotherapy interventions each year and will have access to the data on the costs incurred to provide them. In contrast, a new drug will likely have been given to a few hundred patients and the patient-level data will not be available to the decision maker. Hence, it is unremarkable to observe that some parameters are more uncertain than others, and by extension, some parameters will make a larger contribution to the total decision uncertainty than others. Given that research budgets, like health budgets, are limited, decision makers may be interested in receiving disaggregated analyses of the VOI, to enable them to prioritise how they invest their limited research resources.

There are two distinct forms of disaggregated value of information: the expected value of partial parameter information (EVPPPI) which calculates the EVPI for a specific parameter in the decision model and the EVSI which calculates the expected value of the additional information provided by a study of a specified scale. The formal calculation of each of these is beyond the scope of this text to the extent that there are no hands-on exercises doing either of these calculations. This is primarily because the volume of simulations required for their calculation entails extensive processing times for a model constructed in Excel with simulations programmed in Visual Basic. However, it is important to understand how to undertake these analyses in principle, and the remainder of this chapter focusses on providing both a conceptual understanding of each of these measures and an introduction to the methods for their calculation.

### 12.6.1 *Expected Value of Perfect Parameter Information*

Understanding the EVPPI is a very small step from understanding the EVPI. We are interested in the difference between the Expected NB based upon the decision with current information and the Expected NB if we were certain of the value of the parameter of interest. We obtain our estimate of the latter by running the simulation for each possible value for the parameter of interest whilst allowing all other parameters to vary using PSA. For each run, we record the maximum NB between the treatments being compared. We then take the mean of these to calculate the Expected NB with certainty for the parameter of interest. The difference between the Expected NB with perfect parameter information and the Expected NB with current

information is the maximum value of additional research on the parameter of interest, the EVPPI.

$$EVPI_{\phi} = E_{\phi} \max_j E_{\{\phi,\psi\}} NB(j, \phi, \psi) - \max_j E_{\theta} NB(j, \theta) \quad 12.4$$

Here,  $\phi$  is the specific parameter we are interested in;  $\psi$  is all the other uncertain parameters in the model, and  $\theta$  is the current knowledge on all parameters in the model.

### **12.6.2 Expected Value of Sample Information**

The EVSI is the reduction in the expected cost of making the wrong decision attributable to the additional information provided by a study of a specified sample size ( $n$ ). We quantify this by simulating possible results of studies of the specified size, updating the existing information for the parameter(s) of interest with the simulated sample data to produce a Posterior distribution and then re-estimating the NB for each treatment and identifying which treatment has the higher Expected NB. This process is repeated, and the average of the Expected NB across each simulation is calculated. The difference in the Expected NB based upon current information and the Expected NB based upon the Posterior distributions is the value of the sample information.

$$EVSI = E_D [( \max_j E_{\theta_i} | D NB(j, \theta_i) ] - \max_j E_{\theta} NB(j, \theta) \quad 12.5$$

where  $\theta$  is the parameter of interest and  $D$  is the simulated additional data.

### **12.7 Exercise: Constructing the Net Benefit Probability Map and Calculating the Value of Perfect Information**

The Markov model exercise is very nearly complete. All that remains is to construct the Net Benefit Curve and the NBPM from the simulation data and then calculate the EVPI. As the net benefit breakeven curve is part of the NBPM, we will deal with the construction of both together. Whilst this chapter covers the major steps in producing the NBPM and calculating the expected value of perfect information, we have also provided worked exercises using Excel Workbooks available from <http://medhealth.leeds.ac.uk/costeffectivenessmodelling>. A total of 3 exercises are provided. Exercise 12A constructs the Net Benefit Break Even Curve; 12B constructs the NBPM; and 12C calculates the EVPI.

Remember that the NBPM records the incremental cumulative NB over the time horizon of the model and the uncertainty around the expectation at each time point. In this model, it is the incremental NB of Treatment B compared to Treatment A that

we are interested in. To construct the NBPM, we need to add variables for the cycle-specific incremental QALYs, incremental costs, NB and cumulative NB. The easiest place to add these variables is at the end of the Markov trace data for Treatment B. Remember to add the half-cycle correction at the end of the incremental cost and incremental QALY data. In order to capture the data for all simulations – which you will need to construct the NBPM – you need to create a copy of the cumulative NB variable. The simplest way to do this is to set the column next to your ‘Cumulative Net Benefit’ column equal to the Cumulative Net Benefit column and label it ‘Copy\_Cun\_NB’. Then make the cell references absolute, i.e. cell reference AS22 becomes \$AS\$22.

You can now create an array reporting all the cumulative NB data from each simulation, but you will need to amend your existing PSA macro to ensure that it captures the additional cells containing the additional results and paste them into the extended results array.

The first step is to copy and paste the Copy\_Cum\_NB data so that it is next to your Expected Costs and Outcome data from your existing probabilistic analysis. This data is probably a single row in the worksheet, so you will want to paste the cumulative data as a row, using the paste-special-transpose facility. Table 12.6 shows the format of our model’s per-period probabilistic outputs and simulation array. The per-period results extend to period 61 and the simulations continue all the way down simulation 10,000. We show the first 5 per-period results and the first 5 simulations only.

The simulation results array contains all the information required to construct the NBPM. The NBPM is constructed from the mean value and deciles of the cumulative NB in each period. The easiest way to construct the NBPM in Excel is to use the percentile and average functions. Table 12.7 shows the way we constructed the NBPM data array in our model. Again, it shows 5 periods, whilst the actual model will have 61 periods.

Once you have a table similar to the one shown in Table 12.7, you should be able to construct the NBPM using the line graph facility within Excel. As always, remember to label your axes and give the chart a title. Given the number of contours plotted, it will be useful to have a legend to allow the reader to interpret what each contour plots. Hopefully, your NBPM looks similar to Fig. 12.8.

### 12.7.1 *Calculating the Expected Value of Perfect Information*

In your model, you should already have a column that calculates which technology was selected, on the basis of NMB in each simulation. We labelled this column ‘Optimal’ in our model. It should take the values ‘Treatment A’, ‘Treatment B’ or ‘Either’.

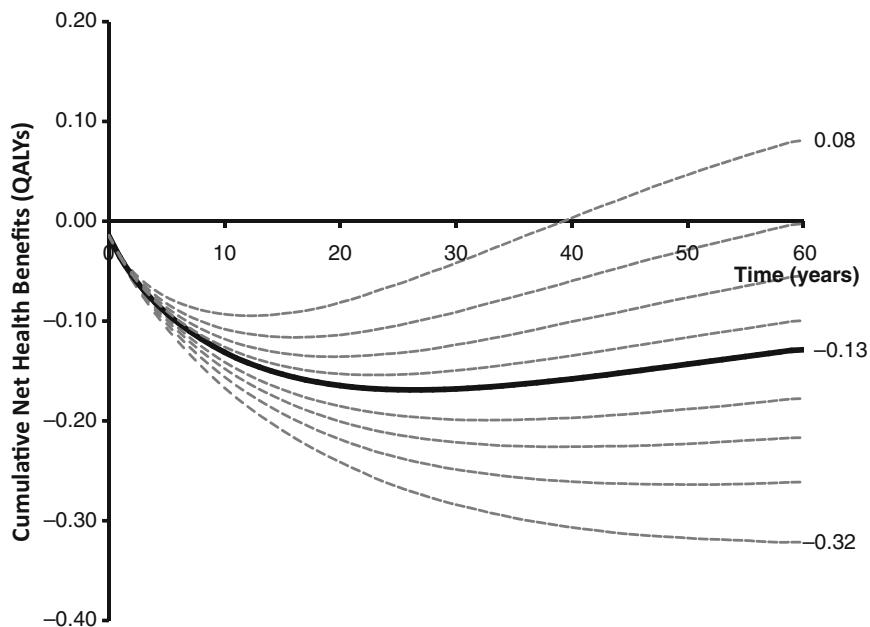
Next, you will need to calculate the NB Loss associated with choosing the intervention based upon the Expected NB over all simulations, rather than being able to choose the best option for each simulation. This is the difference between the NMB if you chose the technology with the highest Expected NMB over all simulations

**Table 12.6** Output format for per-period probabilistic outputs

<b>Simulation</b>	<b>txA_cost</b>	<b>txA_QALY</b>	<b>txB_cost</b>	<b>txB_QALY</b>	<b>txB_cumu_NMB_1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
					\$15,468	0.506	0.837	\$30,478	\$15,468
					<b>txB_cumu_NMB</b>				
					<b>Period</b>				
1	\$16,878	0.405	\$30,665	0.596					
2	\$17,893	0.404	\$30,424	0.640					
3	\$16,931	0.532	\$28,499	0.699					
4	\$15,646	1.307	\$29,998	1.524					
5	\$17,258	-0.609	\$31,527	0.120					

**Table 12.7** Data array for creating net benefit probability map

Deciles and Mean of the Cumulative NHB Per Period	
Period	
0	=PERCENTILE(BF9:BF10008, 0.1)
<i>1st decile</i>	=PERCENTILE(BG9:BG10008, 0.1)
<i>2nd decile</i>	=PERCENTILE(BF9:BF10008, 0.2)
<i>3rd decile</i>	=PERCENTILE(BF9:BF10008, 0.3)
<i>4th decile</i>	=PERCENTILE(BF9:BF10008, 0.4)
<i>Mean</i>	=AVERAGE(BF9:BF10008)
<i>6th decile</i>	=PERCENTILE(BF9:BF10008, 0.6)
<i>7th decile</i>	=PERCENTILE(BF9:BF10008, 0.7)
<i>8th decile</i>	=PERCENTILE(BF9:BF10008, 0.8)
<i>9th decile</i>	=PERCENTILE(BF9:BF10008, 0.9)
	3
	=PERCENTILE(BH9:BH10008, 0.1)
	=PERCENTILE(BH9:BH10008, 0.2)
	=PERCENTILE(BH9:BH10008, 0.3)
	=PERCENTILE(BH9:BH10008, 0.4)
	=AVERAGE(BH9:BH10008)
	=PERCENTILE(BH9:BH10008, 0.6)
	=PERCENTILE(BH9:BH10008, 0.7)
	=PERCENTILE(BH9:BH10008, 0.8)
	=PERCENTILE(BH9:BH10008, 0.9)



**Fig. 12.8** Net benefit probability map with expected breakeven curve

and the Expected NB when you choose the technology with the higher NMB in each simulation. You will need to use a nested IF statement that compares the chosen technology based upon the simulation-specific NB with the technology chosen on the basis of the Expected NB over all the simulations. The EVPI is the Expected NB Loss over all the simulations. Table 12.8 shows the first five simulation results from our model and the formulas we have used to calculate the EVPI. Note that the formula for the EVPI refers to the full 10,000 simulation results. You should find that the EVPI is in the region of \$3,757 per person.

## 12.8 Summary

- Decision uncertainty is the risk of making the wrong decision about a technology: adopting a technology that is not good value or failing to adopt a technology that is good value.
- Health-care reimbursement authorities are increasingly interested in understanding the decision uncertainty and taking this into account in their deliberations.
- CEACs and CEAFFs provide a highly aggregated account of decision uncertainty.
- Examining the distribution of uncertainty regarding the expected costs and benefits over time provides more granular information to decision makers.

**Table 12.8** Calculating the expected value of perfect information

	<b>txA_NB</b>	<b>tkB_NB</b>	Selected	<b>EVPI</b>
	\$8,337	\$4,435	tx A	=AVERAGE (BS9: BS10008)
Probability Cost Effective	=COUNTIF(\$BR9:\$BR10008, "tx A")/COUNT(BK\$9:BK\$ 10008)	0.4183		
	incr_cost	incr_QALY	ICER	
	S5,280	0.046	\$114,620	
Simulation	<b>txA_NB</b>	<b>txB_NB</b>	Optimal	NB Loss
1	\$31,099	\$166	tX A	= IF(\$BR\$2="tx A", MAX(0, BQ9- BP9), IF(\$BRS2 ="tx B", MAX(0, BP9-BQ9), (MAX(BP9, BQ9) - MIN(BP9, BQ9)))
2	-\$19,943	-\$23,286	tx A	\$0
3	-\$19,872	-\$17,512	tx B	\$2,360
4	\$7,942	-\$7,442	tx A	\$0
5	\$23,846	\$17,552	tx A	\$0

- The NBPM is a summary output from probabilistic cost effectiveness models that show how costs and benefits, and the uncertainty around them, are distributed over the time horizon of the analysis.
- NBPMs can help decision makers understand the impact of access with evidence development schemes on value, time to break even and decision uncertainty.
- VOI analyses can allow decision makers to consider the relative value of different access with evidence development schemes compared to delaying reimbursement to allow further research.

## References

- Arrow KJ (1963) Uncertainty and the welfare economics of medical care. *Am Econ Rev* 53: 941–973
- Claxton K, Neumann P, Araki S, Weinstein M (2001) Bayesian value of information analysis: an application to a policy model of Alzheimer's disease. *Int J Technol Assess Health Care* 17(1): 38–55
- Claxton K, Sculpher M, Drummond M (2002) A rational framework for decision making by the National Institute for Clinical Excellence. *Lancet* 360(9334):711–715
- Department of Health (2002) Cost effective provision of disease modifying therapies for people with multiple sclerosis. *Health Service Circular* 2002/004, Stationery Office, London
- Edlin R, Hall P, Wallner K, McCabe C (2014) Sharing risk between payer and provider by leasing health technologies: an affordable and effective reimbursement strategy for innovative technologies. *Value Health* 17(4):438–444

- McCabe CJ, Stafinski T, Edlin R, Menon D (2010) Access with evidence development schemes: a framework for description and evaluation. *Pharmacoeconomics* 28(2):143–152
- McCabe C, Edlin R, Hall P (2013) Navigating time and uncertainty in health technology appraisal: would a map help? *Pharmacoeconomics* 31(9):731–737
- Stafinski T, McCabe CJ, Menon D (2010) Funding the unfundable. Mechanisms for managing uncertainty in decisions on the introduction of new and innovative technologies into healthcare systems. *Pharmacoeconomics* 28(2):113–142
- Walker S, Claxton K, Sculpher M, Palmer S (2012) Coverage with evidence development, only in research, risk sharing or patient access schemes? A framework for coverage decisions. *Value Health* 15(3):570–579

# **Chapter 13**

## **Value of Information in Health Technology Regulation and Reimbursement**

**Abstract** The objective of medical research from a societal or health-care perspective should be to improve the health of a population. In the preceding chapters it is apparent that we need to think about more than simply the effectiveness of the intervention undergoing evaluation if we are to meet this objective. Research is expensive and subjects patients to the risk of experimentation. Both of these factors, set in the context of uncertainty, have the potential to incur opportunity costs within a population. This concluding chapter builds on Chap. 12 setting the context for the use of Value of Information in research prioritisation and research design.

### **13.1 Introduction**

The objective of medical research from a societal or health-care perspective should be to improve the health of a population. In the preceding chapters, it is apparent that we need to think about more than simply the effectiveness of the intervention undergoing evaluation if we are to meet this objective. Research is expensive and subjects patients to the risk of experimentation. Both of these factors, set in the context of uncertainty, have the potential to incur opportunity costs within a population. Opportunity cost may be incurred through the use of a clinically suboptimal treatment strategy, or it may be incurred by suboptimal expenditure that denies more effective treatment to other patients. Where research takes place, there will also be expenditure on the research process and administration; similar expenditure elsewhere in the health system could potentially provide greater health gains.

There is a broad-published literature on the current methods for the design of isolated research studies and clinical trials. In general, the literature focuses on methods for measuring clinically meaningful differences to a specific level of statistical certainty. There is less published guidance on methods for the design of whole research programmes or for prioritising between and within topics for study. Current public research prioritisation mechanisms are far from transparent but mainly rely on predefined criteria which are open to interpretation within peer review and panel discussion procedures. An explicit and reproducible framework that complements this process is required to estimate the potential population-level benefit expected from specific research, including the benefit of reducing reimbursement decision

uncertainty. Decision analysis and Value of Information (VOI) analysis offer a framework for prioritising both within and between research programmes. In this chapter we explore the potential of VOI analysis to inform the design of research, when the aim of the research is to inform subsequent reimbursement decisions. Section 13.2 considers the use of VOI analysis for research prioritisation. Section 13.3 discusses the potential contribution of VOI analyses to research design, and Sect. 13.4 considers the barriers to the greater utilisation of the methodology in the design of clinical research for regulatory and reimbursement purposes.

## 13.2 Value of Information Analysis for Research Prioritisation

Chapter 12 outlined the principles of VOI analysis. The use of VOI analysis has been described as a means of quantifying the decision uncertainty, or expected opportunity cost, if a decision is made to adopt an intervention. It is fairly intuitive to see how the expected value of perfect information (EVPI) provides an estimate of the burden of uncertainty on a decision maker for a single, defined decision problem. If a decision maker is faced with a number of competing decisions for a population, then the EVPI can be used to compare the magnitude of decision uncertainty for each. Where the decision maker has the capacity to invest in further research, they can use the EVPI as a ranking tool for research prioritisation between decision problems.

If decision makers wishes to move beyond the ranking and prioritisation that the EVPI offers, then they need to consider the expected value of sample information (EVSI). EVSI has been described as a means to determine the extent to which the decision uncertainty will be reduced by research of a given design. It can be used in the same way as the EVPI to rank between research designs and is a more robust measure for this purpose. It should, however, be noted that the EVSI is a measure of the potential reduction in the burden of decision uncertainty at the point in time when the decision has to be made. It is not a comprehensive measure of the value of investing in research to reduce that uncertainty. Additional factors need to be taken into account when valuing research and will be outlined in the following sections.

## 13.3 Value of Information Analysis for Research Design

Research has consequences in addition to the consumption of resources invested. Research can be time-consuming and will result in a delay to a final or revised adoption decision. The consequences of this may either be the denial of an effective therapy to patients or the suboptimal investment in an inappropriate therapy whilst the research takes place. Three components that contribute to the value of investing

in research are therefore missing from the EVSI formulation described in Chap. 12: (1) the time that it takes for research to report, (2) the uncertainty around the time it will take research to report and (3) the costs incurred and outcomes experienced by patients whilst research is undertaken and once it has reported. For clarity in this book, we refer to the measure that includes these elements as the expected net present value of sample information (ENPVSI).

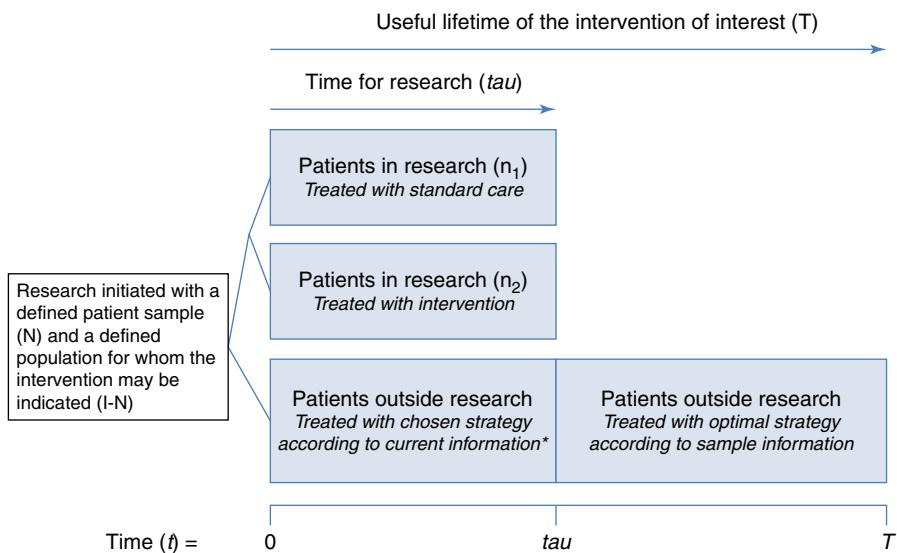
### ***13.3.1 Calculating the Expected Net Present Value of Sample Information***

Information will be forthcoming from research that will report at a future point in time,  $t$ . As before, the research provides data  $X_{\theta_t}$  that will update parameter values from the prior  $\theta$  to the posterior  $\theta | X_{\theta_t}$ .

The total cost of the research consists of not only the cost of undertaking the research, but also the value of the health expected to be gained or foregone by patients participating *within* the research. The costs and health gain or decrement incurred to patients *outside* the research study also need to be taken into account. If they receive standard care during the research when a new intervention is expected to be superior, then they will suffer a net loss between the time of the decision to invest in the research and the time when the research reports and leads to an updated decision. Both the costs of the research and the benefits from the research need to be discounted to reflect the fact that the new information and any change in the reimbursement decision will occur at an uncertain point in the future. Therefore, to accurately capture the ENPVSI, it is necessary to model:

1. The Expected Costs and Health Outcomes of patients involved in the study up to the point of the research reporting
2. The Expected Costs and Health Outcomes of all patients not involved in the research study up to the point of the research reporting
3. The Expected Costs and Health Outcomes of all patients, both those involved in the study, and those outside the study, after the research has reported
4. The uncertain timing of the reporting of the research

Steps one to three involve sampling from the Expected Net Benefit (NB) in the usual way. An intuitive method by which to implement step four is to incorporate a trial simulation model into the EVSI calculation to represent the uncertain estimate for the time for research to take place ( $\tau$ ). Tau is likely to depend on a number of uncertain factors including the expected effectiveness of each strategy and the baseline rate of the event of interest where clinical outcomes are time dependent, the uncertain time to set up research and an uncertain recruitment rate. It may also incorporate a risk that research will not complete at all. This would then allow us to expand the modelling of the payoff from research to take account of the costs and outcomes for patients within and outwith the research up to the time when the research reports and the reimbursement decision reviewed.



**Fig. 13.1** Diagrammatic presentation of calculation of Net Benefit from sample information (Hall et al. 2012).\* This will depend on whether the strategy with the highest NB is used or whether this is constrained by an OIR or OWR arrangement

If we consider a research design comparing standard care with a single intervention in a two-arm randomised controlled trial, the assessment of NB over time needs to consider the per-patient NB in each of four groups of patients, multiplied by the number of patients over the relevant time period for each group. The ENPVI will be the combination of the Expected NBs of these groups of patients (Fig. 13.1):

1. Within research:
  - (a) Treated with standard care ( $\text{popNB}_{\text{trial.1}}$ )
  - (b) Treated with intervention ( $\text{popNB}_{\text{trial.2}}$ )
2. Outwith research ( $\text{popEVSI}_{\text{out}}$ )
  - (a) Treated with standard care
  - (b) Treated with intervention

Given that we are proposing research, the Expected NB for patients who are outwith the research should rely on the expectation given the proposed research. By contrast, the Expected NB for the patients in the trial is read out from the simulated trial result on the basis of the notional sample size. Constraints are placed on those patients for whom the disease is incident prior to the time research reports such that the choice of treatment may be constrained by inclusion in one arm of the trial. If the patient is not in the trial, the choice of optimal treatment will be based on current information – this may be subject to the specific decision rules employed by the decision maker at time zero.

The expected net present value of sample information therefore is given by

$$\text{ENPVSI} = \sum_{j=1}^2 \text{popNB}_{\text{trial},j} + \text{popNB}_{\text{out}} - \text{popNB}_{\text{current}}$$

When comparing  $J$  treatment strategies, this generalises to

$$\text{ENPVSI} = \sum_{j=1}^J \text{popNB}_{\text{trial},j} + \text{popNB}_{\text{out}} - \text{popNB}_{\text{current}}$$

where assuming that  $\theta I$  and  $\theta Ic$  are independent

$$\begin{aligned} \text{popNB}_{\text{trial},j} &= E_{\tau} \left\{ E_{X_{\theta_I}} \left[ E_{\theta_I^c, \theta_I | X_{\theta_I}} \text{NB}(j; \theta_I, \theta_I^c) \right] \cdot \sum_{t=1}^{\tau} \frac{n_{jt}}{(1+r)^t} \right\} \\ \text{popNB}_{\text{out}} &= E_{\tau} \left\{ E_{X_{\theta_I}} \left[ \max_j E_{\theta_I^c, \theta_I | X_{\theta_I}} \text{NB}(j, \theta_I, \theta_I^c) \right] \cdot \left[ \sum_{t=1}^{\tau} \frac{I_t - n_t}{(1+r)^t} + \sum_{t=\tau+1}^T \frac{I_t}{(1+r)^t} \right] \right\} \\ \text{popNB}_{\text{current}} &= \max_j E_{\theta} \text{NB}(j, \theta) \cdot \sum_{t=1}^T \frac{I_t}{(1+r)^t} \end{aligned}$$

where

$nt$ =number of patients in trial during time interval  $t$

$\tau$ = time for further sampling (time for research to report)

$T$ =time over which decision is pertinent

The Monte Carlo sampling algorithm for implementation of ENPVSI through a two-level simulation is outlined in Appendix 13.1.

If we wish to consider the special case where approval is given for a new intervention Only in Research (OIR) or the intervention is adopted conditional on research taking place (Only with Research – OWR), then the same framework can be applied with additional constraints applied at time zero. Specifically, the NB attributable to patients treated outside research prior to  $\tau$  is constrained to either standard care in the case of OIR or intervention in the case of OWR.

## 13.4 Is Decision Theory Ready to Inform Trial Design?

VOI methods based on decision modelling continue to be a focus of intense development by health economists and Bayesian statisticians. However, there are a number of challenges that remain to be overcome if their potential is to be realised more widely in the clinical research community.

### ***13.4.1 Structuring a Decision Problem***

As with any decision modelling, it is essential to adequately represent the underpinning clinical pathway. All aspects of a problem need to be addressed without creating an overcomplex model, which increases the risk of analytical error. Adequately representing structural uncertainty remains difficult and omission of a key parameter can have major knock-on consequences in VOI analyses, leading in turn to error in attributing priorities for research. This is a problem for any type of model-based economic evaluation, not just for research prioritisation. Sensitivity analysis around alternative model structures is the current best solution.

### ***13.4.2 Evidence Synthesis and Model Parameterisation***

Evidence synthesis and model parameterisation include, for example, the challenges of eliciting prior information from experts to inform model parameters where empirical research has not yet taken place (see Chap. 2). There is also a requirement to consider correlation between different parameters of a model and to accurately extrapolate short-term data over a relevant time horizon. Reliance on surrogate outcomes and the transferability of data between health-care settings also pose challenges.

### ***13.4.3 Computational and Statistical Challenges***

EVPI and EVSI can be calculated by a variety of methods depending on the validity of assumptions about model linearity and normality (Ades et al. 2004; Brennan et al. 2007; Hall et al. 2012; Strong and Oakley 2013; Strong et al. 2014). Such methods are easier to implement than those described in this chapter, but their use with the complex fully non-parametric models often required for health technology assessment, limits their usefulness. Non-parametric EVSI calculation poses a much higher computational burden, and although analytical developments are ongoing, much more work is needed. Currently the best solution is through the use of high-performance computing hardware.

### ***13.4.4 Adoption by Regulatory Organisations and Reimbursement Agencies***

It is the reimbursement authorities and other regulators who have the power to ensure methods are implemented, perhaps with cooperation from licensing authorities. It is also their responsibility to ensure that adequate analytic standards are adhered to. A number of challenges must be overcome before they can do this.

These include the problems associated with reversing a decision after the emergence of further information, problems with conducting further research after an intervention is adopted as routine care and the potential incentive for jurisdictions to wait for others to conduct research (free-rider problems).

### ***13.4.5 Adoption by Public Research Commissioners and Clinical Trialists***

The framework for economic evaluation in health care that has emerged over the last 30 years is unfamiliar to much of the clinical research community. There remains mistrust of decision modelling which can seem opaque to those without technical expertise (Claxton et al. 2005). Traditional methods for creating hypotheses and designing trials have developed over many decades. Clinicians and statisticians have learnt to work together to enhance the internal validity of clinical research: adoption of an alternative framework unsurprisingly faces some resistance. To be effective, clinicians, trialists and health economists must work together to deliver this alternative paradigm that acknowledges the reality of opportunity cost. Realistic and understandable applied examples are needed to demonstrate to a relevant audience how the methods described can improve the design of important clinical trials. Whilst these are starting to emerge, more are needed to get to the tipping point whereby the design of research processes targets the efficient production of information required to inform regulation and reimbursement decisions.

### ***13.4.6 Industrial Development of Health Technologies***

Prioritising or designing research on the basis of societal benefit might seem irrelevant in the context of industry-funded drug research, when the objective is profit maximisation rather than societal health benefit. However, ensuring that pharmaceutical research adequately informs public reimbursement decision makers is likely to work to the advantage of companies seeking to speed up market access. The provision of clear goal posts in this respect is essential in order to provide the industrial research designer a clear incentive to generate unambiguous cost effectiveness-based outcomes. Indeed, as regulators move towards formally demanding cost effectiveness evidence, there will be a need for companies to improve the efficiency of their research programmes in meeting these endpoints, in addition to purely clinical outcomes. Recently, the use of VOI analysis has been proposed as part of the Value Engineered Translation (VET) framework (Bubela and McCabe 2014), which triages translation investment technologies according to their potential to meet value-based market access criteria and then identifies the key evidence-based investments covering manufacturing and regulatory concerns as

well as conventional safety and effectiveness issues. The VET framework is being used by the developers in relation to personalised medicine, stem cell and oncology therapies in the pre-Phase III clinical development space.

### **13.5 Value of Information in the Evolving Regulatory and Reimbursement Environments**

An explicit framework for research priority in health-care setting is long overdue. The widespread adoption of decision analytic models for health technology reimbursement processes means that decision theoretic methods are an increasingly appropriate framework. They allow estimation of the value of conducting research and can help ensure that sufficient evidence is generated for adequately informed reimbursement decisions at, or close to, the time of licensing. They also support the design of post market research that is coherent with the premarket research activities, thus enabling developers and public authorities to make informed decisions about which evidence to require prior to patient access to technologies and which evidence can be efficiently gathered after this point. Given the increasing interest in accelerated approval of new technologies by regulatory authorities, this capacity for the coherent design of pre- and post market research should be useful to both regulators and reimbursement authorities (Drug Administration 2012; EMEA 2006; Health 2014). The VOI framework can provide a shared methodological framework for regulatory and reimbursement authorities discussions regarding on post market access research requirements. It also allows these authorities to demonstrate that the cost of such research is justified in relation to the value of the evidence that the required research will provide. Further, it can promote a consistent approach to the evaluation of technologies across clinical indications and between regulatory processes for different types of technology. Such consistency of methods and processes can reduce the uncertainty associated with investing in developing new technologies, which should be commercially attractive to the developer and manufacturers' communities.

In the coming years it should be possible for health-care payers, regulators and reimbursement authorities, working with clinical trialists, statisticians and health economists, to establish in advance how much and what type of evidence is required to inform a health technology adoption decision, thereby improving the likelihood that effective therapies will be available for their patients, at prices that ensure that the new technologies have a positive impact on population health.

### **13.6 Summary**

- Research has consequences in that it can be time-consuming and will result in a delay to a final or revised adoption decision which may lead to either be the denial of an effective therapy to patients or the suboptimal investment in an inappropriate therapy whilst the research takes place

- ENPVI allows the analyst to take into account the time that it takes for research to report, the uncertainty around the time it will take research to report and the costs incurred and outcomes experienced by patients whilst research is undertaken and once it has reported
- To accurately capture EVPSI the analyst needs to model the Expected Costs and Health Outcomes of patients involved in the study up to the point of the research reporting; the Expected Costs and Health Outcomes of all patients not involved in the research study up to the point of the research reporting; the Expected Costs and Health Outcomes of all patients, both those involved in the study and those outside the study after the research has reported; and the uncertain timing of the reporting of the research.
- VOI methods still pose challenges including adequate representation of structural uncertainty, how evidence is synthesised and computational burden.

## Appendix

### *Appendix 13.1: General Monte Carlo Sampling Algorithm for Calculation of Population ENPVI*

Adapted from Ades et al. (2004)

$\theta_I$ =parameters of interest (here assumed independent of  $\theta_I^c$ )

First record the net benefit of an optimal decision based on current information. Then define a proposed piece of research from which data  $X_{\theta_I}$  will be collected to inform  $\theta_I$ .

A1. For  $i = 1, 2, \dots, N$  simulations

- B1. Draw a sample  $\theta_I^{(i)}$  from the prior (baseline) distribution of  $\theta_I$ .
- B2. Draw a sample  $X_{\theta_I^{(i)}}$  from the distribution of the sufficient statistic  $X_{\theta_I}|\theta_I^{(i)}$  arising from a new study of defined size.
- B3. Calculate posterior (updated) expected net benefits for each strategy j, using an inner Monte Carlo simulation loop using the posterior distribution  $\theta_I^{(i)}|X_{\theta_I^{(i)}}$ .
- B4. Calculate expected net benefits for each strategy  $j$  given the likelihood  $X_{\theta_I^{(i)}}$ , evaluated at its mean, using an inner Monte Carlo simulation loop.
- B5. Find the strategy  $j$  maximising expected net benefit for simulation  $i$  based on B3.
- B6. Draw a sample from the distribution of time to trial reporting ( $\tau$ ) using  $X_{\theta_I^{(i)}}$ .
- B7. Using the expected net benefit given the mean of the likelihood  $X_{\theta_I^{(i)}}$  (B4.), allocate net benefit to patients allocated to trial arms for each strategy  $j$  for each time interval up to  $\tau$ , discounted.
- B8. Using the posterior expected net benefits (B3.), record the population net benefit for patients not in trial for time intervals prior to time  $\tau$  who receive the optimal strategy  $j$  given a decision based on the prior expected net benefits up to time  $\tau$ , discounted.

- B9. Record the population net benefit for the optimal strategy  $j$  given a decision based on the posterior expected net benefits using the discounted population for each time interval after the trial has reported.
- B10. Record the sum of the expected net benefits over all groups in B7, B8 and B9.
- A2. Find the average of the population expected net benefits (B10), over the  $N$  simulations. This is the population expected value of a decision based on sample information.
- A3. Subtract from this the population expected value of a decision based on current information to give the ENPVI.

## References

- Ades AE, Lu G, Claxton K (2004) Expected value of sample information calculations in medical decision modeling. *Med Decis Making* 24(2):207–227
- Brennan A, Kharroubi S, O'Hagan A, Chilcott J (2007) Calculating partial expected value of perfect information via Monte Carlo sampling algorithms. *Med Decis Making* 27(4):448–470
- Bubela T, McCabe C (2014) Value engineered translation: developing biotherapeutics for health system needs. *Evidence Based Diabetes Management* 20(10L):1
- Claxton K, Egginton S, Ginnelly L, Griffin S, McCabe C, Philips Z, Tappenden P, Wailoo A (2005) A pilot study of value of information analysis to support research recommendations for the National Institute for Health and Clinical Excellence. CHE Research Paper. [http://www.york.ac.uk/media/che/documents/papers/researchpapers/rp4\\_Pilot\\_study\\_of\\_value\\_of\\_information\\_analysis.pdf](http://www.york.ac.uk/media/che/documents/papers/researchpapers/rp4_Pilot_study_of_value_of_information_analysis.pdf)
- European Medicines Agency (2006) Guideline for the procedure for accelerated assessment pursuant to article 14(9) of Regulation (EC) No 726/2004. [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Regulatory\\_and\\_procedural\\_guideline/2009/10/WC500004136.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2009/10/WC500004136.pdf)
- Food and Drug Administration (2012) FDA safety and innovation act. <http://www.fda.gov/RegulatoryInformation/Legislation/FederalFoodDrugandCosmeticActFDCAAct/SignificantAmendmentstotheFDCAAct/FDASIA/default.htm>
- Hall PS, Edlin R, Kharroubi S, Gregory W, McCabe C (2012) Expected net present value of sample: from burden to investment. *Med Decis Making* 32(3):E11–E21
- Health Canada (2014) Protection of Canadians from unsafe drugs act (Vanessa's Law). <http://www.hc-sc.gc.ca/dhp-mps/legislation/unsafedrugs-droguesdangereuses-eng.php>
- Strong M, Oakley JE (2013) An efficient method for computing single parameter partial expected value of perfect information. *Med Decis Making* 33(6):755–766
- Strong M, Oakley JE, Brennan A (2014) Estimating multiparameter partial expected value of perfect information from a probabilistic sensitivity analysis: a non-parametric regression approach. *Med Decis Making* 34(3):311–326